

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 15:40:52 ; Search time 88 Seconds
(without alignments)
1242.757 Million cell updates/sec

Title: US-10-052-664-1

Perfect score: 3597

Sequence: 1 MAPWPELGDAQPNPKYLEG.....SREAGYVPSDSKTECTAL 689

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 13Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3597	100.0	689	21	AA18613
2	3594	99.9	689	22	ABB50299
3	3589	99.8	689	23	ABG96383
4	3581.5	99.6	690	22	AAW78631
5	3581.5	99.6	700	22	ABB12311
6	3581.5	99.6	700	22	AAW79615
7	3570.5	99.3	690	19	AAW77414
8	3240	90.1	631	22	AAW93865
9	2098	58.3	397	22	AAW93860

10	1529	42.5	599	22	AAE06591	Human protein havi
11	742	20.6	155	19	AAW77415	Human sodium depen
12	418.5	11.6	388	22	AAQ92763	C Glutamicum prote
13	410	11.4	411	22	ABW76765	Corynebacterium gl
14	280.5	7.8	549	23	ABF29511	Streptococcus poly
15	274.5	7.6	544	23	ABB47553	Listeria monocytog
16	263.5	7.3	543	24	ABP81331	Streptococcus pneu
17	263.5	7.3	543	24	ABU00859	S. pneumoniae type
18	183.5	5.1	120	19	AAW77597	Possible sodium-de
19	142.5	4.0	441	23	ABB47592	Listeria monocytog
20	133.5	3.7	443	23	ABB48414	Listeria monocytog
21	131.5	3.7	90	22	ABG27784	Novel human diagno
22	130.5	3.6	74	22	ABG23596	Novel human diagno
23	130.5	3.6	74	22	ABG23597	Novel human diagno
24	130.5	3.6	74	22	ABG23601	Novel human diagno
25	129.5	3.6	886	23	AAE16785	Human transporter
26	128.5	3.6	524	22	AAU54281	Propionibacterium
27	128.5	3.6	556	24	ABB82976	Human SLC22A relat
28	128.5	3.6	565	22	ABG11289	Novel human diagno
29	128.5	3.6	565	22	ABG11291	Novel human diagno
30	127.5	3.5	701	23	AAE24209	Rat obesity-specif
31	126.5	3.5	307	24	ABP81373	Streptococcus pneu
32	126.5	3.5	307	24	ABU01250	S. pneumoniae type
33	126.5	3.5	517	23	ABF66239	Bifidobacterium lo
34	126	3.5	398	22	ABG19008	Novel human diagno
35	125.5	3.5	1743	24	ABU07906	Novel human secret
36	125	3.5	547	23	ABG93114	S. cerevisiae BAX-
37	124.5	3.5	456	24	ABJ19067	Pathogen specific
38	122	3.4	122	22	ABG27074	Novel human diagno
39	122	3.4	130	22	ABG23598	Novel human diagno
40	122	3.4	130	22	ABG27785	Human transporter
41	121.5	3.4	509	23	ABB83901	Human transporter
42	121.5	3.4	509	23	ABG61543	Drosophila melanog
43	121.5	3.4	896	22	ABG63778	Novel human diagno
44	121	3.4	89	22	ABG27446	Streptococcus poly
45	120	3.3	728	23	ABP26246	

ALIGNMENTS

RESULT 1
AA18613
ID AA18613 standard; Protein; 689 AA.
XX
AC AA18613;
XX
DT 15-JAN-2001 (first entry)
XX
DE A human sodium phosphate cotransporter, designated Npt2B.
XX
KW Human; sodium phosphate cotransporter; Npt2B; intestinal epithelial cell;
KW phosphate absorption; inorganic phosphate metabolism; phosphatemia;
KW osteomalacia; rickets; hyperparathyroidism; renal disease.
XX
OS Homo sapiens.
XX
PN AU200014979-A.
XX
PD 31-AUG-2000.
XX
PF 09-FEB-2000; 2000AU-0014979.
XX
PR 09-FEB-1999; 99US-0119321.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Cannon PD, Sankurati S;
XX
DR WPI; 2000-579513/55.
XX
DR N-PSDB; AAA75484.
XX
PT Human sodium-phosphate cotransporter Npt2B, useful e.g. for identifying

PT agents for treating hyper- and hypo-phosphatemia, and related nucleic
acid -
XX
BS Claim 3; Fig 1; 42pp; English.
XX
CC The present sequence represents a human sodium phosphate cotransporter,
designated Npt2B. It is expressed on the apical surface of intestinal
CC epithelial cells. Npt2B is a human type II protein, which is responsible
CC for absorption and uptake of inorganic phosphate in the intestine. The
CC protein is used to identify specific agonists and antagonists, and to
CC raise antibodies. The agonists and antagonists are potentially useful
CC for treating disease of inorganic phosphate metabolism, especially
CC hyper- and hypo-phosphatemia, e.g. osteomalacia, rickets,
CC hyperparathyroidism and renal disease.
XX
SQ Sequence 689 AA;

Query Match 100.0%; Score 3597; DB 21; Length 689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPPELGDAPNPDKYLEGAGOOPTAPDKSKETKNTNTEAPVTKIELLPSTATLID 60
DB 1 MAPPELGDAPNPDKYLEGAGOOPTAPDKSKETKNTNTEAPVTKIELLPSTATLID 60

QY 61 EPTVEDDPNWLPTLODSGKIKWSEKRTKGLKILCFQIGRLLILLGLFFVFCSLDILSSA 120
DB 61 EPTVEDDPNWLPTLODSGKIKWSEKRTKGLKILCFQIGRLLILLGLFFVFCSLDILSSA 120

QY 121 FOLVGGKMGAGOFFNSSIMSNPLGLVIGLVTLVQSSSTSTSIWSSVSSLLTVRAA 180
DB 121 FOLVGGKMGAGOFFNSSIMSNPLGLVIGLVTLVQSSSTSTSIWSSVSSLLTVRAA 180

QY 181 IPIIMGANIGTSTNTIIVAMQVGRSEPRPRAAGTVDHDFNNLSVLVLLPEVATHYL 240
DB 181 IPIIMGANIGTSTNTIIVAMQVGRSEPRPRAAGTVDHDFNNLSVLVLLPEVATHYL 240

QY 241 EIITQLIVBSFHFHFKNGEDAPLLKVIKPTFKLIVOLDKKVTSQIAMDDEKAKNKSIVKI 300
DB 241 EIITQLIVBSFHFHFKNGEDAPLLKVIKPTFKLIVOLDKKVTSQIAMDDEKAKNKSIVKI 300

QY 301 WKTFNKTQINVTVPSTANTCPSLCTWDGIQNTWKNTVYKENIAKQCHI FWNPELPD 360
DB 301 WKTFNKTQINVTVPSTANTCPSLCTWDGIQNTWKNTVYKENIAKQCHI FWNPELPD 360

QY 361 LAVGTILLLSLVLCGCLIMVKILGSLVKGQAVVIKKTINTDPPFPFAMLTGVIALIL 420
DB 361 LAVGTILLLSLVLCGCLIMVKILGSLVKGQAVVIKKTINTDPPFPFAMLTGVIALIL 420

QY 421 VGAGMTFVQSSSVFTSALTPLIGIVITTIERAYPLTLGNSNIGTTTTLAALASPQNAL 480
DB 421 VGAGMTFVQSSSVFTSALTPLIGIVITTIERAYPLTLGNSNIGTTTTLAALASPQNAL 480

QY 481 RSSLIQIALCHFPNFTSGILLWYPPFRLPIRMAKGLNISKYKFWFAVFLIIFFLP 540
DB 481 RSSLIQIALCHFPNFTSGILLWYPPFRLPIRMAKGLNISKYKFWFAVFLIIFFLP 540

QY 541 LTVFGLSLAGRVLVGVGVVVFVFIILVLCRLQSCRPVLPKQLQNNWFLPLWRSLSK 600
DB 541 LTVFGLSLAGRVLVGVGVVVFVFIILVLCRLQSCRPVLPKQLQNNWFLPLWRSLSK 600

QY 601 PWDVAVSKFTGFQMRCCCRVCCACCLLCGCPKCCCKCEDLEAQBEGQDVPVKA 660
DB 601 PWDVAVSKFTGFQMRCCCRVCCACCLLCGCPKCCCKCEDLEAQBEGQDVPVKA 660

QY 661 PETPDNITISREAGQVPAQSKTECTAL 689
DB 661 PETPDNITISREAGQVPAQSKTECTAL 689

RESULT 2

ABB50299

ID, ABB50299 standard; Protein; 689 AA.

XX
AC ABB50299;
XX
DT 08-FEB-2002 (first entry)
XX
DE HOST-5 (Nap1-11b) ovarian tumour marker protein, SEQ ID NO:147.
XX
KW Ovarian tumour marker gene; human; overexpression; upregulation;
KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
KW identification; serous cystadenoma; borderline serous tumour;
KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
KW immune response pathway; cell proliferation regulation; protein folding;
KW membrane localised; secreted; therapeutic target; cytostatic;
KW gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200175177-A2.
XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US10947.
XX
PR 03-APR-2000; 2000US-194336P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
XX
PI WPI, 2001-626450/72.
XX
DR N-PSDB; ABA83184.
XX
PT Detecting and identifying ovarian tumor, identifying increased risk for
PT developing ovarian cancer, and determining effectiveness of ovarian
PT cancer treatment, by measuring expression level of ovarian tumor marker
PT gene -
XX
PS Claim 24; Page 139-140; 140pp; English.
XX
CC The invention relates to methods for diagnosing and prognosing ovarian
CC tumours in an individual via the detection and measurement of the
CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
CC ABA83181 and ABA83183). The methods of the invention are useful for
CC detecting an ovarian tumour in a patient, for identifying an individual
CC at increased risk for developing ovarian cancer, in prognostic tests for
CC assessing the relative severity of ovarian cancer, in tests for
CC monitoring a patient in remission from ovarian cancer and in tests for
CC monitoring disease status in a patient being treated for ovarian cancer.
CC The methods can additionally be used to identify a particular tumour as
CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
CC mucinous cystadenoma, borderline mucinous tumour, mucinous
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
CC tumour. The ovarian tumour marker genes of the invention were identified
CC using SAGE (serial analysis of gene expression) and were found to be
CC overexpressed in a broad variety of ovarian epithelial tumour cells
CC relative to normal ovarian epithelial cells. The marker genes are
CC implicated in immune response pathways, in the regulation of cell
CC proliferation and in protein folding, and many of these are membrane-
CC localised or secreted. In addition to their use as diagnostic and
CC prognostic markers, the ovarian tumour marker genes or their encoded
CC proteins may be used as therapeutic targets for their treatment and
CC prevention of ovarian cancer. Sequences ABB50257-ABB50299 represent
CC proteins encoded by ovarian tumour marker genes of the invention.
XX
SQ Sequence 689 AA;

Query Match 99.9%; Score 3594; DB 22; Length 689;

Best Local Similarity 99.9%; Pred. No. 0;		Matches 688; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MAPWPELDAQNPDPKYLEGAGQQTAPDKSKETNKNTAPVTKIELLPSTATLID	60
Db	1	MAPWPELDAQNPDPKYLEGAGQQTAPDKSKETNKNTAPVTKIELLPSTATLID	60
QY	61	EPTVEDDPNNLPTLQDSGKWSERDTKGKILCFQGGIGRLILLGLFYFFVCSLDLSSA	120
Db	61	EPTVEDDPNNLPTLQDSGKWSERDTKGKILCFQGGIGRLILLGLFYFFVCSLDLSSA	120
QY	121	FQLVGKMGAGOFFNSNSIMSNPLGLVIGLVTLVQSSSTSTSVVSMVSSLLVRAA	180
Db	121	FQLVGKMGAGOFFNSNSIMSNPLGLVIGLVTLVQSSSTSTSVVSMVSSLLVRAA	180
QY	181	IPILMGANTGTSITNTIIVLMQVDRSEFRAPAGATVHDFNWLVLVPEVATHYL	240
Db	181	IPILMGANTGTSITNTIIVLMQVDRSEFRAPAGATVHDFNWLVLVPEVATHYL	240
QY	241	EIITQILVSEFFKNGEDAPDLKAVITKPTKLIIVQLODKVISOIAMNDEKAKNKLVI	300
Db	241	EIITQILVSEFFKNGEDAPDLKAVITKPTKLIIVQLODKVISOIAMNDEKAKNKLVI	300
QY	301	WCKTFTNKQINVTVPSTANTCPSLCWTDGIONTMKNVYKXENIAKQHIIFNPHLPD	360
Db	301	WCKTFTNKQINVTVPSTANTCPSLCWTDGIONTMKNVYKXENIAKQHIIFNPHLPD	360
QY	361	LAVGTILLILSLVLCGCLIMVKILGSLVKGQVATVTKTINTDPPPPAMLTGYLATL	420
Db	361	LAVGTILLILSLVLCGCLIMVKILGSLVKGQVATVTKTINTDPPPPAMLTGYLATL	420
QY	421	VGAGMTFIVQSSSVFTSALTPLIGVITITERRAYPLTIGSNIGTITTTAILAALASFGNAL	480
Db	421	VGAGMTFIVQSSSVFTSALTPLIGVITITERRAYPLTIGSNIGTITTTAILAALASFGNAL	480
QY	481	RSSLQIALCHFFPNISGILLWYPPTRLPPIRMAGKIGNISAKYRWFAVYLIIPFLIP	540
Db	481	RSSLQIALCHFFPNISGILLWYPPTRLPPIRMAGKIGNISAKYRWFAVYLIIPFLIP	540
QY	541	LTVPGLSLAGVAVGVGVVVFVFIILVLCRLQSCRPVLPKLQNNWFLPWRSUK	600
Db	541	LTVPGLSLAGVAVGVGVVVFVFIILVLCRLQSCRPVLPKLQNNWFLPWRSUK	600
QY	601	PDAVSVKFTGCFQMRCCCCCRVCCRACCLLCCPKCCCKCCEDLEAQGGQDVPVKA	660
Db	601	PDAVSVKFTGCFQMRCCCCCRVCCRACCLLCCPKCCCKCCEDLEAQGGQDVPVKA	660
QY	661	PETFDNITISREAQGEVPASDSKTECTAL	699
Db	661	PETFDNITISREAQGEVPASDSKTECTAL	699
RESULT 3			
ABG96383			
ID	ABG96383 standard; Protein; 689 AA.		
XX	AC ABG96383;		
XX	DT 11-DEC-2002 (first entry)		
DE	Human ovarian cancer marker M463.		
XX	Human; ovarian cancer; marker; cancer; familial history; brain disorder;		
KW	Central nervous system disorder; bacterial meningitis; viral meningitis;		
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;		
KW	brain herniation; inflammation; encephalitis; testicular disorder;		
KW	nontuberculous granulomatous orchitis; connective tissue disorder;		
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;		
XX	histological type; carcinogenic; ovarian cancer marker.		
OS	Homo sapiens.		
XX	WO200271928-A2.		
PN			

XX	19-SEP-2002.	
PD		
XX	14-MAR-2002; 2002WO-US07826.	
XX	14-MAR-2001; 2001US-276025P.	
PR	14-MAR-2001; 2001US-276026P.	
PR	10-AUG-2001; 2001US-311732P.	
PR	19-SEP-2001; 2001US-323580P.	
PR	26-SEP-2001; 2001US-324967P.	
PR	26-SEP-2001; 2001US-325102P.	
PR	26-SEP-2001; 2001US-325149P.	
XX	(MILL-) MILLENNIUM PHARM INC.	
PA		
PI	Monahan JE, Ganavarrapu M, Hoersch S, Kamatkar S, Kovatis SG;	
PI	Meyers RE, Morrissey MP, Olandt PU, Sen A, Vieby PO, Mills GB;	
PI	Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;	
XX	WPI; 2002-723277/78.	
DR	N-PSDB; ABS76479.	
XX	Assessing whether a patient is afflicted with ovarian cancer, useful in	
PT	assessing the stage or progression of the disease, comprises comparing	
PT	the expression level of a cancer marker in a sample from a patient and	
PT	from a non cancer patient -	
XX	Disclosure; Page 349-351; 481pp; English.	
PS		
XX		
CC	The present invention relates to a new method for assessing whether a	
CC	patient is afflicted with ovarian cancer. The method involves comparing	
CC	the expression level of a marker in a patient sample and the normal level	
CC	of expression of the marker in a control non-ovarian cancer sample, where	
CC	the marker is selected from 363 cancer markers described in the	
CC	specification. The method of the invention is useful in diagnosing or	
CC	characterising cancer, in detecting the presence of cancer as early as	
CC	possible, and the recurrence of ovarian cancer. The method may also be of	
CC	particular use with patients having an enhanced risk of developing	
CC	ovarian cancer (e.g. patients having a familial history of ovarian	
CC	cancer). The cancer markers may be used in the management and treatment	
CC	of e.g. brain and central nervous system disorders (e.g. bacterial and	
CC	viral meningitis, Alzheimer's disease or Parkinson's disease), brain	
CC	inflammations (e.g. cerebral oedema, hydrocephalus or brain herniations),	
CC	testicular disorders (e.g. bacterial or viral meningitis or encephalitis),	
CC	connective tissue disorders, or heart disorders (e.g. ischaemic heart	
CC	disease or atherosclerosis). The compositions and methods may also be	
CC	used in assessing the histological type of neoplasm associated with	
CC	ovarian cancer, monitoring the progression of ovarian cancer,	
CC	determining whether ovarian cancer has metastasized or is likely to	
CC	metastasize, selecting a composition for inhibiting ovarian cancer,	
CC	assessing the ovarian carcinogenic potential of a compound, or	
CC	inhibiting ovarian cancer or at risk of developing ovarian cancer. The	
CC	present amino acid sequence represents one of the ovarian cancer markers	
CC	described in the invention.	
XX		
SQ	Sequence 689 AA;	
Query Match 99.8%; Score 3589; DB 23; Length 689;		
Best Local Similarity 99.7%; Pred. No. 0;		
Matches 687; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
QY	1	MAPWPELDAQNPDPKYLEGAGQQTAPDKSKETNKNTAPVTKIELLPSTATLID 60
Db	1	MAPWPELDAQNPDPKYLEGAGQQTAPDKSKETNKNTAPVTKIELLPSTATLID 60
QY	61	EPTVEDDPNNLPTLQDSGKWSERDTKGKILCFQGGIGRLILLGLFYFFVCSLDLSSA 120
Db	61	EPTVEDDPNNLPTLQDSGKWSERDTKGKILCFQGGIGRLILLGLFYFFVCSLDLSSA 120
QY	121	FQLVGKMGAGOFFNSNSIMSNPLGLVIGLVTLVQSSSTSTSVVSMVSSLLVRAA 180
Db	121	FQLVGKMGAGOFFNSNSIMSNPLGLVIGLVTLVQSSSTSTSVVSMVSSLLVRAA 180

QY 181 IPIIMGANIGTSITNTIIVAMOVGDRSEFRRAFAGATVHDFNNLSVLVLLPVEVATHYL 240
 Db 181 IPIIMGANIGTSITNTIIVAMOVGDRSEFRRAFAGATVHDFNNLSVLVLLPVEVATHYL 240
 QY 241 EIITQIIVSFHFHFKNGEDAPDLKVIITKPTFKLIVQLDKKVISQIAMNDEKAKNKSIVKI 300
 Db 241 EIITQIIVSFHFHFKNGEDAPDLKVIITKPTFKLIVQLDKKVISQIAMNDEKAKNKSIVKI 300
 QY 301 WCKTFTNKTOINVTVPSTANCTSPSLCWTGDIQNWTKMKNVTYKNIACQHI FYNFHLDP 360
 Db 301 WCKTFTNKTOINVTVPSTANCTSPSLCWTGDIQNWTKMKNVTYKNIACQHI FYNFHPD 360
 QY 361 LAVGFTILLISLLVLCGLIMIVKILGSLVKGQVATVTKTINTDPPPPFAWLGYLAIL 420
 Db 361 LAVGFTILLISLLVLCGLIMIVKILGSLVKGQVATVTKTINTDPPPPFAWLGYLAIL 420
 QY 421 VQAGMTFVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTTTTALALASPQNAL 480
 Db 421 VQAGMTFVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTTTTALALASPQNAL 480
 QY 481 RSSLQIALCHFFFNISGILLWYPIPTPLIRMAKGLGNISAKYRFAVYLIIFPFLIP 540
 Db 481 RSSLQIALCHFFFNISGILLWYPIPTPLIRMAKGLGNISAKYRFAVYLIIFPFLIP 540
 QY 541 LTVFGLSLAGNRVLVGVGVVVFIIILVLCRLQLQSCRPVLPKQLQNNWFLPLMWSLK 600
 Db 541 LTVFGLSLAGNRVLVGVGVVVFIIILVLCRLQLQSCRPVLPKQLQNNWFLPLMWSLK 600
 QY 601 PWDVSKFTGCFQMRCCCCCVRACCLLCGPKFCRCCKCCEBDEAQBQGVVKA 660
 Db 601 PWDVSKFTGCFQMRCCCCCVRACCLLCGPKFCRCCKCCEBDEAQBQGVVKA 660
 QY 661 PETFDNITISREAGVPSADSKTECTAL 689
 Db 661 PETFDNITISREAGVPSADSKTECTAL 689
 RESULT 4
 AAM78631
 ID AAM78631 standard; Protein; 690 AA.
 XX AAM78631;
 AC AAM78631;
 XX AAM78631;
 DT 06-NOV-2001 (first entry)
 XX Human protein SEQ ID NO 1293.
 DE Human
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200157190-A2.
 XX 09-AUG-2001.
 PD 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US04098.
 PF 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663361.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 P4

PI Zhao QA, Wang D, Zhang J, Zhang J, Ren F, Chen R, Wang ZH;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR N-PSDB; AAK51764.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 20; Page 3546-3548; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 690 AA;
 SQ
 Query Match 99.6%; Score 3581.5; DB 22; Length 690;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 688; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MAPPELGDQPNPKYLEGAGGOPTAPDKSKETNK-NNTEAPVTKEILLPSVSTATLI 59
 Db 1 MAPPELGDQPNPKYLEGAGGOPTAPDKSKETNKNDTEAPVTKEILLPSVSTATLI 60
 QY 60 DEPTVEDDPNLTQDSGIKWSERDTKGKILCFQGIQGRILLGLFLYFVCSLDILSS 119
 Db 61 DEPTVEDDPNLTQDSGIKWSERDTKGKILCFQGIQGRILLGLFLYFVCSLDILSS 120
 QY 120 AFQLVGNKAGOFFNSNSINPLLLGLVITVLVWSSSTSTSI VVWVSSSLTVRA 179
 Db 121 AFQLVGNKAGOFFNSNSINPLLLGLVITVLVWSSSTSTSI VVWVSSSLTVRA 180
 QY 180 AIPILMGNIGTSITNTIIVAMOVGDRSEFRRAFAGATVHDFNNLSVLVLLPVEVATHY 239
 Db 181 AIPILMGNIGTSITNTIIVAMOVGDRSEFRRAFAGATVHDFNNLSVLVLLPVEVATHY 240
 QY 240 LEITQIIVSFHFHFKNGEDAPDLKVIITKPTFKLIVQLDKKVISQIAMNDEKAKNKSIVK 299
 Db 241 LEITQIIVSFHFHFKNGEDAPDLKVIITKPTFKLIVQLDKKVISQIAMNDEKAKNKSIVK 300
 QY 300 IWCKTFTNKTOINVTVPSTANCTSPSLCWTGDIQNWTKMKNVTYKNIACQHI FYNFHLP 359
 Db 301 IWCKTFTNKTOINVTVPSTANCTSPSLCWTGDIQNWTKMKNVTYKNIACQHI FYNFHLP 360
 QY 360 DLAVGTILLISLLVLCGLIMIVKILGSLVKGQVATVTKTINTDPPPPFAWLGYLAI 419
 Db 361 DLAVGTILLISLLVLCGLIMIVKILGSLVKGQVATVTKTINTDPPPPFAWLGYLAI 420
 QY 420 LVGAGMTFVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTTTTALALASPQNA 479
 Db 421 LVGAGMTFVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTTTTALALASPQNA 480
 QY 480 LRSSLQIALCHFFFNISGILLWYPIPTPLIRMAKGLGNISAKYRFAVYLIIFPFLI 539
 Db 481 LRSSLQIALCHFFFNISGILLWYPIPTPLIRMAKGLGNISAKYRFAVYLIIFPFLI 540
 QY 540 PLTVFGLSLAGNRVLVGVGVVVFIIILVLCRLQLQSCRPVLPKQLQNNWFLPLMWSL 599
 Db 541 PLTVFGLSLAGNRVLVGVGVVVFIIILVLCRLQLQSCRPVLPKQLQNNWFLPLMWSL 600
 QY 600 KPMDAVVSKFTGCFQMRCCCCCVRACCLLCGPKFCRCCKCCEBDEAQBQGVVVK 659

Db 601 KPDAVSKFTGCEQWCCCRVCRACCLCGCPKCCCKCELEBAQEGQDVPVK 660
Qy 660 APETFDNITISRAQGEVPASDSKTECTAL 689
Db 661 APETFDNITISRAQGEVPASDSKTECTAL 690
RESULT 5
ID ABB12311 standard; peptide; 700 AA.
AC ABB12311;
XX 11-JAN-2002 (first entry)
XX Human Na-dependent phosphate transporter homologue, SEQ ID NO:2681.
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
XX cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
XX antifungal; vulnery; antiulcer.
XX Homo sapiens.
XX OS
XX WO200157188-A2.
XX PN
XX 09-AUG-2001.
XX PD
XX 05-FEB-2001; 2001WO-US03800.
XX PF
XX 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PT
XX (HYSB-) HYSEQ INC.
XX PA
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457740/49.
XX DR N-PSDB; ABA09555.
XX DR
XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
XX PT
XX Claim 20; Page 329; 1963pp; English.
XX PS
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.

CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
XX Sequence 700 AA;
SQ
Query Match 99.6%; Score 3581.5; DB 22; Length 700;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 688; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 MAMPBGLDAQNPDKYLEGAGQQTAPDKSKETNK-NNTTEAPVTKIELPSYSTATLI 59
DB 11 MAMPBGLDAQNPDKYLEGAGQQTAPDKSKETNKTDTEAPVTKIELPSYSTATLI 70
QY 60 DEPTVEDDPPNLTLDQSGIKWSEKRTKGIKLCFFQGIKILLLGLLGLFFVCSLDILSS 119
DB 71 DEPTVEDDPPNLTLDQSGIKWSEKRTKGIKLCFFQGIKILLLGLLGLFFVCSLDILSS 130
QY 120 AFQIVGKAGQOFFSNGSINSPNLLGLVIGLVTVLVVQSSSTSTSIIVSNVSSLLTVRA 179
DB 131 AFQIVGKAGQOFFSNGSINSPNLLGLVIGLVTVLVVQSSSTSTSIIVSNVSSLLTVRA 190
QY 180 AIPITMGANIGTSINTTIVALMOVGDRSFRAPAGATVHDFFNWLVLVLLPVEVATHY 239
DB 191 AIPITMGANIGTSINTTIVALMOVGDRSFRAPAGATVHDFFNWLVLVLLPVEVATHY 250
QY 240 LEITITLIVESPHFKNGEDAPDLLKVIITKPTKLIVOLDKKVISOIAMNDEKAKNKLIVK 299
DB 251 LEITITLIVESPHFKNGEDAPDLLKVIITKPTKLIVOLDKKVISOIAMNDEKAKNKLIVK 310
QY 300 IWCKTFNKTQINVTVPSTANCTSPSLCWTGDIQNWTKNVTYKENTAKOHLFVNFHLP 359
DB 311 IWCKTFNKTQINVTVPSTANCTSPSLCWTGDIQNWTKNVTYKENTAKOHLFVNFHLP 370
QY 360 DLAVGTILLILSLVLCGLIMIVKILGSLVLAGQVATVIKKTINTDPPPPFAWLGYLAI 419
DB 371 DLAVGTILLILSLVLCGLIMIVKILGSLVLAGQVATVIKKTINTDPPPPFAWLGYLAI 430
QY 420 LVGAGMTFVQSSSVFTSALTPLIGITVITIERAIVPLTIGSNIGTITTTTALAALASPNA 479
DB 431 LVGAGMTFVQSSSVFTSALTPLIGITVITIERAIVPLTIGSNIGTITTTTALAALASPNA 490
QY 480 LRSLOIALCHFFFNISGILLWYPIPTPLPRMAKAGNISAQYRWFVYLIIFFFLI 539
DB 491 LRSLOIALCHFFFNISGILLWYPIPTPLPRMAKAGNISAQYRWFVYLIIFFFLI 550
QY 540 PLTVFGLSLAGRWLVGVGVVVFVFIILVLCRLQLQSCRPVLPKLIQNNVFLPMWESL 599
DB 551 PLTVFGLSLAGRWLVGVGVVVFVFIILVLCRLQLQSCRPVLPKLIQNNVFLPMWESL 610
QY 600 KPWDVAVSKFTGCEQWCCCRVCRACCLCGCPKCCCKCELEBAQEGQDVPVK 659
DB 611 KPWDVAVSKFTGCEQWCCCRVCRACCLCGCPKCCCKCELEBAQEGQDVPVK 670
QY 660 APETFDNITISRAQGEVPASDSKTECTAL 689

Db 671 APETFDNITISREAOGEVPASDSKTECTAL 700

RESULT 6
AAW79615
ID AAW79615 standard; Protein; 700 AA.
XX AC AAW79615;
XX DT 06-NOV-2001 (first entry)
XX DE Human protein SEQ ID NO 3261.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200157190-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US04098.
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663561.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WI; 2001-476283/51.
XX DR N-PSDB; AAK52748.
XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX FS Claim 20; Page 296-297; 6221pp; English.
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAW80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX SQ Sequence 700 AA;
Query Match 99.6%; Score 3581.5; DB 22; Length 700;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 688; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 MAPWPELGDAQPNPDKYLGAAGQOPTAPDKSKETNK-NNTAPVTKIELLPYSTATLI 59
D, 11 MAPWPELGDAQPNPDKYLGAAGQOPTAPDKSKETNKNTDTEAPVTKIELLPYSTATLI 70

QY 60 DEPTVDDPNNLPTLQDSGKWKSERDTKGKILCFQGIQGRILILLGLFLYFFVCSLDLSS 119
DB 71 DEPTVDDPNNLPTLQDSGKWKSERDTKGKILCFQGIQGRILILLGLFLYFFVCSLDLSS 130
QY 120 AFOLVGKMGAGOFFSNSIMSNPLGLVIGLVATVTLVLOSSTSTSVVSVSSSLTVA 179
DB 131 AFOLVGKMGAGOFFSNSIMSNPLGLVIGLVATVTLVLOSSTSTSVVSVSSSLTVA 190
QY 180 AIPIMGANIGTSITNTIVALMQVGRDRFRFAFAGATVHDFFNWLSVLVLEVEVATH 239
DB 191 AIPIMGANIGTSITNTIVALMQVGRDRFRFAFAGATVHDFFNWLSVLVLEVEVATH 250
QY 240 LEIITOLIVESHFKNGEDAPDLKVIITPFTKLIIVOLDKVIISQIAMNDEKAKNLSLVK 299
DB 251 LEIITOLIVESHFKNGEDAPDLKVIITPFTKLIIVOLDKVIISQIAMNDEKAKNLSLVK 310
QY 300 IWCKTFNKTQINVTVPSTANCTSPSLCWTGDIQNWTKMNTYKENTAKOHIHFNHFLP 359
DB 311 IWCKTFNKTQINVTVPSTANCTSPSLCWTGDIQNWTKMNTYKENTAKOHIHFNHFLP 370
QY 360 DLAVGTILLISLLVLCGLIMIVKILGSLVKQVATVIKKTINTDPPFPFALTCYLAI 419
DB 371 DLAVGTILLISLLVLCGLIMIVKILGSLVKQVATVIKKTINTDPPFPFALTCYLAI 430
QY 420 LVGAGMTFVQSSSVFTSALTPLIGITVITIRAYPHLTGSGNIGTTTITAILAALASPGNA 479
DB 431 LVGAGMTFVQSSSVFTSALTPLIGITVITIRAYPHLTGSGNIGTTTITAILAALASPGNA 490
QY 480 LRSLOIALCHEFFNIGSILMWPIPTPLPIRMAKGLGNISAKYRFAVYLIIFFFLI 539
DB 491 LRSLOIALCHEFFNIGSILMWPIPTPLPIRMAKGLGNISAKYRFAVYLIIFFFLI 550
QY 540 PLTVFGLSLAGRWLVGVGVVVFIIILVLCRLQSRCPVLPKLNWNFLPLMWSL 599
DB 551 PLTVFGLSLAGRWLVGVGVVVFIIILVLCRLQSRCPVLPKLNWNFLPLMWSL 610
QY 600 KPWDVAVSKFTGCFOMRCCCVCCVCCACCLCGCPKCCCKCEDLEAORGDVPUK 659
DB 611 KPWDVAVSKFTGCFOMRCCCVCCVCCACCLCGCPKCCCKCEDLEAORGDVPUK 670
QY 660 APETFDNITISREAOGEVPASDSKTECTAL 689
DB 671 APETFDNITISREAOGEVPASDSKTECTAL 700

RESULT 7.
AAW77414
ID AAW77414 standard; Protein; 690 AA.
XX AC AAW77414;
XX DT 02-FEB-1999 (first entry)
XX DE Human sodium dependent phosphate transporter IPT-1.
XX DE IPT-1; sodium dependent phosphate transporter; human;
XX kidney failure; kidney disease; uraemic bone disease; cancer;
XX diagnosis; therapy; vaccine.
XX OS Homo sapiens.
XX PN EP875569-A1.
XX PD 04-NOV-1998.
XX PF 09-APR-1998; 98EP-0302815.
XX PR 23-SEP-1997; 97US-0935433.
XX PR 28-APR-1997; 97US-0044974.
XX PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Feild J;

XX WPI; 1998-559435/43.

XX N-PSDB; AAV77414.

XX New DNA encoding sodium-dependent phosphate transporter protein
 PT IPT-1 - used to prevent, treat and diagnose e.g. chronic renal
 PT failure, end-stage renal disease, uraemic bone disease and cancer
 XX

XX Claim 1; Page 17-19; 24pp; English.

XX This is the amino acid sequence of human sodium dependent phosphate
 CC transporter IPT-1, as deduced from the nucleotide sequence of a
 CC cDNA clone (see AAV59498). It shows about 76% identity in 691 amino
 CC acids with bovine sodium dependent phosphate transporter and 55%
 CC identity in 625 amino acids with human Napi-3. Methods are
 CC provided for the production of IPT-1 polypeptides in recombinant
 CC host cells. Such polypeptides can be used to raise antibodies and
 CC also in a claimed method for identifying compounds which inhibit or
 CC agonise the IPT-1 polypeptide. Agonists can be used to treat a
 CC subject in need of enhanced activity or expression of IPT-1, while
 CC antagonists can be used to treat a subject having need to inhibit
 CC activity or expression of IPT-1. A claimed process for diagnosing
 CC a disease or susceptibility to disease related to IPT-1 expression
 CC or activity involves analysing for the presence or amount of IPT-1
 CC expression in a sample. IPT-1 polypeptides and polynucleotides can
 CC be used for prevention, treatment and diagnosis of e.g. chronic
 CC renal failure, end-stage renal disease, uraemic bone disease and
 CC cancer.

XX Sequence 690 AA;

Query Match 99.3%; Score 3570.5; DB 19; Length 690;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 687; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MAPWELGDAQPNPKYLEGAGGQPTAPDKSKETNK-NNTEAPVTKIELLPSTATLI 59
 DB 1 MAPWELGDAQPNPKYLEGAGGQPTAPDKSKETNKNTDTEAPVTKIELLPSTATLI 60
 QY 60 DEPTVEDDNNPTLQDSGIKWSERDTKGIKLCFFQIGRLILILGTYFFVCSLDLSS 119
 DB 61 DEPTVEDDNNPTLQDSGIKWSERDTKGIKLCFFQIGRLILILGTYFFVCSLDLSS 120
 QY 120 AFQLVCKWAGOFFNSNSIMNPLGLVIGLVTVLVQSSSTSTSVVSWSSSLTVRA 179
 DB 121 AFQLVCKWAGOFFNSNSIMNPLGLVIGLVTVLVQSSSTSTSVVSWSSSLTVRA 180
 QY 180 AIPFINGANIGTSINTTIVALMQVDSRPRAFAGATVHDFFNWLSVLVLPVEVATHY 239
 DB 181 AIPFINGANIGTSINTTIVALMQVDSRPRAFAGATVHDFFNWLSVLVLPVEVATHY 240
 QY 240 LEIITOLIVESHFKNGEDAPDLKLVTKPTKLIIVOLDKKVISOIAMDKEKAKNSLYK 299
 DB 241 LEIITOLIVESHFKNGEDAPDLKLVTKPTKLIIVOLDKKVISOIAMDKEKAKNSLYK 300
 QY 300 IWCKTFTNTKQINVTVPSTANTCSPSLCWDGIGQNTWMTYKENTAKQHLFVNFLHP 359
 DB 301 IWCKTFTNTKQINVTVPSTANTCSPSLCWDGIGQNTWMTYKENTAKQHLFVNFLHP 360
 QY 360 DLAVGTILILSLVLCGLIMIVKILGSLVKQVATVTKTINTDPPFPFAMLTGYLAI 419
 DB 361 DLAVGTILILSLVLCGLIMIVKILGSLVKQVATVTKTINTDPPFPFAMLTGYLAI 420
 QY 420 LVGAGMTFIVQSSSVFTSALTPLIGIVITIERAYPLTIGSNIGTITTAIALASPNA 479
 DB 421 LVGAGMTFIVQSSSVFTSALTPLIGIVITIERAYPLTIGSNIGTITTAIALASPNA 480
 QY 480 LRSLQIALCHFFFNISGILLWYIPTRPIRMAGLGNISAKYRFAVYFYLIIFFFLI 539
 DB 481 LRSLQIALCHFFFNISGILLWYIPTRPIRMAGLGNISAKYRFAVYFYLIIFFFLI 540
 QY 540 PLTVFGLSLAGWRVLVGVGVVVFIIIVLCLELQSCRPVLPKQLQNNWFLPLMWSRL 599

DB 541 PLTVFGLSLAGWRVLVGVGVVVFIIIVLCLELQSCRPVLPKQLQNNWFLPLMWSRL 600
 QY 600 KPWDVAVSKFTGCFQMRCCGCCVCCRACCLLCGCPKCCRCCKCEDLBAAGQGVPK 659
 DB 601 KPWDVAVSKFTGCFQMRCCGCCVCCRACCLLCGCPKCCRCCKCEDLBAAGQGVPK 660
 QY 660 APETFDNITISREAQGEVPASDSKTECTAL 689
 DB 661 APETFDNITISREAQGEVPASDSKTECTAL 690

RESULT 8

AAM93865

ID AAM93865 standard; Protein; 631 AA.

XX AAM93865;

AC AAM93865;

XX 06-NOV-2001 (first entry)

DT Human polypeptide, SEQ ID NO: 3967.

XX Human; full length cDNA; cDNA synthesis; oligo-capping.

XX Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99JP-0194486.

XX 11-JAN-2000; 2000JP-0118774.

XX 02-MAY-2000; 2000JP-0183785.

XX (HELI-) HELIX RES INST.

XX Ota T. Nishikawa T. Isogai T. Hayashi K. Ishii S. Kawai Y;

XX Wakamatsu A. Sugiyama T. Nagai K. Kojima S. Otsuki T. Koga H;

XX WPI; 2001-52455/58.

XX N-PSDB; AAK94823.

XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -

XX Claim 8; SEQ ID NO 3967; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesising the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 631 AA;

Query Match 90.1%; Score 3240; DB 22; Length 631;

Best Local Similarity 91.2%; Pred. No. 1.3e-301;

Matches 629; Conservative 1; Mismatches 0; Indels 60; Gaps 2;

QY 1 MAPWELGDAQPNPKYLEGAGGQPTAPDKSKETNK-NNTEAPVTKIELLPSTATLI 59

DB 1 MAPWELGDAQPNPKYLEGAGGQPTAPDKSKETNKNTDTEAPVTKIELLPSTATLI 60

QY 60 DEPTVEDDNNPTLQDSGIKWSERDTKGIKLCFFQIGRLILILGTYFFVCSLDLSS 119

Db 61 DEPTVEDDPNLPTLQDSGKWSERDTKGKILCFQGIKRLILLGLFLYPPVCSLDILSS 120
 QY 120 AFQLVGKMGAGOFFNSSTMSNPLGLVIGVATVVLVQSSSTSTSIIVSMVSSSLITVRA 179
 Db 121 AFQLVGKMGAGOFFNSSTMSNPLGLVIGVATVVLVQSSSTSTSIIVSMVSSSLITVRA 180
 QY 180 AIPIMGANIGTSITNTIIVAMQVCDRSEPRRAFAGATVHDFNMLSVLVLPEVATHV 239
 Db 181 AIPIMGANIGTSITNTIIVAMQVCDRSEPRRAFAGATVHDFNMLSVLVLPEVATHV 240
 QY 240 LEITQILVESHFHNGEDAPDLKIVTKPTKLIIVOLDKKVISOIAMNDEKAKNSLVK 299
 Db 241 LEITQILVESHFHNGEDAPDLKIVTKPTKLIIVOLDKKVISOIAMNDEKAKNSLVK 300
 QY 300 IWCKTFNKTQINVTVPSTACTSPSLCWTGDTQNTMKNVTYKENIAKQHFVNPHLP 359
 Db 301 IWCKTFNKTQINVTVPSTACTSPSLCWTGDTQNTMKNVTYKENIAKQHFVNPHLP 360
 QY 360 DLAVGTILLILSLVLCGLIMVKILGSLVKGQVATVIKKTINTDPPFPFAMLTGYLAI 419
 Db 361 DLAVGTILLILSLVLCGLIMVKILGSLVKGQVATVIKKTINTDPPFPFAMLTGYLAI 420
 QY 420 LVGAGMTFIVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTTTAILAALASFGNA 479
 Db 421 LVGAGMTFIVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTTTAILAALASFGNA 480
 QY 480 LRSLQLALCHFFNFISGILLWYPIPTPLPIRMKAGLGNISAKYRWFVAYFLIIPFLLI 539
 Db 481 LRSLQLALCHFFNFISGILLWYPIPTPLPIRMKAGLGNISAKYRWFVAYFLIIPFLLI 540
 QY 540 PLTVFGLSLAGWVVLGVGVVVFIIILVLCRLQSRCPVLPKLLQNNWFLPLMWSL 599
 Db 541 PLTVFGLSLAGWVVLGVGVVVFIIILVLCRLQSRCPVLPKLLQNNWFLPLMWSL 541
 QY 600 KPMDAVVSKFTGCFQMRCCCRVCCACCLCGCKCCCKCEDLEBAQEGQDVPVK 659
 Db 601 KPMDAVVSKFTGCFQMRCCCRVCCACCLCGCKCCCKCEDLEBAQEGQDVPVK 601
 QY 660 APETFDNITISREAQGEVPSDSKTECTAL 689
 Db 631 APETFDNITISREAQGEVPSDSKTECTAL 631

RESULT 9

AAW93860
 ID AAW93860 standard; Protein; 397 AA.

XX AC AAW93860;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide, SEQ ID NO: 3955.

XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.

XX PN EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-0114089.

XX PR 08-JUL-1999; 99JP-0194486.

XX PR 11-JAN-2000; 2000JP-0118774.

XX PR 02-MAY-2000; 2000JP-0183765.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX FI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX XX

DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94816.
 XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX Claim 8; SEQ ID NO 3955; 1380pp + sequence listing; English.
 XX The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesising the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX Sequence 397 AA;

Query Match 58.3%; Score 2098; DB 22; Length 397;

Best Local Similarity 98.3%; Pred. No. 2.8e-192;

Matches 396; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 287 MNDEKAKNSLVKIWKCTFTNKTQINVTVPSTACTSPSLCWTGDTQNTMKNVTYKENI 346

Db 1 MNDEKAKNSLVKIWKCTFTNKTQINVTVPSTACTSPSLCWTGDTQNTMKNVTYKENI 60

QY 347 AKQCHIFVNFHPLDPAVGTILLILSLVLCGLIMVKILGSLVKGQVATVIKKTINTDP 406

Db 61 AKQCHIFVNFHPLDPAVGTILLILSLVLCGLIMVKILGSLVKGQVATVIKKTINTDP 120

QY 407 PFPFAMLTGYLAILVAGMTFIVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTT 466

Db 121 PFPFAMLTGYLAILVAGMTFIVQSSSVFTSALTPLIGIGVITIERAYPLTGLSGNIGTTT 180

QY 467 TAILAALASFGNALRSSLOIALCHFFNFISGILLWYPIPTPLPIRMKAGLGNISAKYRW 526

Db 181 TAILAALASFGNALRSSLOIALCHFFNFISGILLWYPIPTPLPIRMKAGLGNISAKYRW 240

QY 527 FAVFYLIIFPFLIPLTVFGLSLAGWVVLGVGVVVFIIILVLCRLQSRCPVLPKLL 586

Db 241 FAVFYLIIFPFLIPLTVFGLSLAGWVVLGVGVVVFIIILVLCRLQSRCPVLPKLL 300

QY 587 QNWNFLPLMWSLKPMDAVVSKFTGCFQMRCCCRVCCACCLCGCKCCCKCED 646

Db 301 QN-----NWSLKPMDAVVSKFTGCFQMRCCCRVCCACCLCGCKCCCKCED 354

QY 647 LEEAQEGQDVPVKAPETFDNITISREAQGEVPSDSKTECTAL 689

Db 355 LEEAQEGQDVPVKAPETFDNITISREAQGEVPSDSKTECTAL 397

RESULT 10

AAE06591

ID AAE06591 standard; Protein; 599 AA.

XX AC AAE06591;

XX DT 25-SEP-2001 (first entry)

XX DE Human protein having hydrophobic domain, HP03878.

XX KW Human; hydrophobic domain; gene therapy; nutritional supplement;

XX KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;

XX KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;

XX KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;

XX KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;

XX KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;

XX KW contraceptive; antiinfertility; antiinflammatory.

CC This is the partial amino acid sequence of a human sodium dependent
CC phosphate transporter IPT-1 variant polypeptide, as deduced from the
CC DNA sequence of a polynucleotide (see AAV59499) of the invention.
CC Methods are provided for the production of IPT-1 polypeptides (see
CC also AAW77414) in recombinant host cells. Such polypeptides can be
CC used to raise antibodies and also a claimed method for identifying
CC compounds which inhibit or agonise the IPT-1 polypeptide. Agonists
CC can be used to treat a subject in need of enhanced activity or
CC expression of IPT-1, while antagonists can be used to treat a
CC subject having need to inhibit activity or expression of IPT-1. A
CC claimed process for diagnosing a disease or susceptibility to
CC disease related to IPT-1 expression or activity involves analysing
CC for the presence or amount of IPT-1 expression in a sample. IPT-1
CC polypeptides and polynucleotides can be used for prevention,
CC treatment and diagnosis of e.g. chronic renal failure, end-stage
CC renal disease, uraemic bone disease and cancer.

XX SQ Sequence 155 AA;

Query Match 20.6%; Score 742; DB 19; Length 155;
Best Local Similarity 98.0%; Pred. No. 9.7e-63;
Matches 145; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 399 KKTINTDPPFFFAWLTGYLAILVAGMTFFVQSSSVFTSALTPLIGIGVITIERAYPLTL 458
Db 1 KKTINTDPPFFFAWLTGYLAILVAGMTFFVQSSSVFTSALTPLIGIGVITIERAYPLTL 60

QY 459 GSNIGTWTITAILAALASPGNALRSSLOALCHFFFNISGILLWYIPTRLPINAKGLG 518
Db 61 GSNIGTWTITAILAALASPGNALRSSLOALCHFFFNISGILLWYIPTRLPINAKGLG 120

QY 519 NISAKYRWFVAVFLLIIFPLIPLTVFGL 546
Db 121 NISAKYRWFVAVFLLIIFPLIPLTVFAL 148

RESULT 12
AAG92763
ID AAG92763 standard; Protein; 388 AA.
XX AAG92763;
XX AAG92763;
XX 26-SEP-2001 (first entry)
XX C glutamicum protein fragment SEQ ID NO: 6517.
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX Corynebacterium glutamicum.
XX EP1108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-0127688.
XX 16-DEC-1999; 99JP-0377484.
XX 07-APR-2000; 2000JP-0159162.
XX 03-AUG-2000; 2000JP-0280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX N-PSDB; AAH67982.
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analysing
XX expression profile or pattern of a gene and identifying homologous gene

XX Claim 17; SEQ ID NO: 6517; 246bp + Sequence Listing; English.
XX PS The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of corynebacterium, measuring expression amount and
XX analysing the expression profile or expression pattern of a gene derived
XX from Corynebacterium, and identifying a homologue of a gene derived
XX from corynebacterium. Corynebacterium bacteria are useful for producing
XX amino acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described
XX in the exemplification of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office.

XX SQ Sequence 388 AA;

Query Match 11.6%; Score 418.5; DB 22; Length 388;
Best Local Similarity 26.9%; Pred. No. 4.3e-31;
Matches 121; Conservative 80; Mismatches 163; Indels 85; Gaps 8;

QY 100 LILLGLFVFFVCSLDLSSAFQLVGCKMAGOFFSNS--SIMSNPLGLGVLTVLVQ 157
Db 13 LLVLLSLVILIGINLLDGVGF-----GTFSTQMYQVAKDPLGLIGILATLVQ 66

QY 158 SSSTSTSVVSVSSSLTVRAAIPIMGANIGTSTINTVVALMOVGDRFRAPAGAT 217
Db 67 SSTTTTLTAVTGIVSVFPAIPIILGANIGTITAMLVAFSYVGRERFKKAFVAA 126

QY 218 VHDFENWLSVLLVPEVATHLEILITQLIVESFHFKNGEDAP--DLLKVTIKPFTKLIV 275
Db 127 MHWFNVLLVLLFVVELLHFFPTTISGATATETITTCGSLTSGVMTKIFDPTQLLG 186

QY 276 QLDKKVISQIAMNDEKAKNKSIVKICKTFTNKTQINVTVPSTANCTSPSLCWTGQNW 335
Db 187 M--NGLIGSIG-----NPSISAIVC----- 204

QY 336 TMRNVTYKENIAKQHI FVNPHLPDLAVGTHILILSLVLOCLIMIVKILGSLVKQVA 395
Db 205 -----LVVGTILILISVRAMS-----SQIRITTAATVT 232

QY 396 TVIKKTINTDPPFPFPAWLTGYLAILVAGMTFFVQSSSVFTSALTPLIGIGVITIERAYP 455
Db 233 SIMDKVINPENSFKATILSNFWSFILGVLFILMTVASSVTVASMQPVAASGVVKQKELLG 292

QY 456 LTLGSNIGTITTAILAALASPGNALRSSLOALCHFFFNISGILLWYIP-FTLPLIRMA 514
Db 293 VILGANVGTITVAMFATFALVSDQGBFAIQALIHILVNFVGTGALLVLCIPQLANVILHA 352

QY 515 KGLQNISAKYRWFVAVFLLIIFPLIPLTV 543
Db 353 EKTANLTARSYSITLTIATVTVLVPSAV 381.

RESULT 13
AAB76765
ID AAB76765 standard; Protein; 411 AA.
XX AAB76765;
XX AAB76765;
XX 11-APR-2001 (first entry)
XX Corynebacterium glutamicum MCT protein SEQ ID NO:512.
XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
XX membrane construction and membrane transport protein; petroleum spill;
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;
XX identification; microorganism; fine chemical production; transformation;
XX genome mapping; genetic engineering.
XX Corynebacterium glutamicum.
OS

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 15:48:07 ; Search time 49 Seconds
(without alignments)
1352.249 Million cell updates/sec

Title: US-10-052-664-1
Perfect score: 3597
Sequence: 1 MAPWELGDAQNPDKYLEG.....SREAGQVPASDKTECTAL 689

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2851	79.3	693	2 S49228	sodium-dependent p
2	1721	47.8	637	2 A48189	sodium/phosphate c
3	1706.5	47.4	853	2 A54366	sodium/phosphate c
4	1698.5	47.2	839	2 B48189	sodium/phosphate c
5	1682	46.8	637	2 A54000	sodium/phosphate c
6	1681.5	46.7	642	2 I46534	renal sodium-depen
7	732.5	20.4	504	2 T27914	hypothetical prote
8	559.5	15.6	382	2 D82295	nptA protein VC067
9	283	7.9	555	2 D89770	conserved hypothet
10	274.5	7.6	544	2 A81178	transport protein
11	263.5	7.3	543	2 E95057	Na/Pi cotransporte
12	262.5	7.3	543	2 G97926	conserved hypothet
13	247.5	6.9	543	2 G83825	hypothetical prote
14	242.5	6.7	569	2 E90554	conserved hypothet
15	218.5	6.1	593	2 H71283	conserved hypothet
16	203	5.6	310	2 F69952	Na+/Pi cotransport
17	194.5	5.4	310	2 A83819	hypothetical prote
18	179	5.0	253	2 AF1535	transport protein
19	165	4.6	543	2 C65209	hypothetical 59.5
20	161	4.5	543	2 B91246	probable alpha hel
21	161	4.5	543	2 H86093	probable alpha hel
22	146.5	4.1	441	2 AB1542	PTS system, Lichen
23	146.5	4.1	682	2 AG3351	transporter EMB107
24	146	4.1	582	2 C82961	conserved hypothet
25	145.5	4.0	543	2 AI1011	probable membrane
26	142.5	4.0	441	2 AD1184	PTS system, Lichen
27	139.5	3.9	550	2 AD2852	Na+/Pi-cotransport
28	139.5	3.9	555	2 B97629	hypothetical prote
29	139.5	3.9	630	2 T38110	major facilitator

30	136	3.8	520	2 H90447	amino acid transpo
31	135.5	3.8	442	2 A83743	magnesium citrate
32	133.5	3.7	443	2 AG1335	hypothetical prote
33	132.5	3.7	539	2 E97025	spoVB related memb
34	131	3.6	164	2 T24272	hypothetical prote
35	129	3.6	152	2 T18975	hypothetical prote
36	129	3.6	654	2 E64245	periplasmic phosph
37	128	3.6	188	2 T15651	hypothetical prote
38	128	3.6	511	2 T11467	cytochrome-c oxida
39	128	3.6	513	2 T13782	cytochrome-c oxida
40	127.5	3.5	524	2 H84141	L-lactate permease
41	126.5	3.5	307	2 B95099	membrane protein
42	126.5	3.5	307	2 A97967	conserved hypothet
43	126.5	3.5	498	2 T14236	NADH2 dehydrogenas
44	125	3.5	547	2 S53920	SNGL protein - yea
45	124.5	3.5	456	2 B90069	hypothetical prote

ALIGNMENTS

RESULT 1

S49228
sodium-dependent phosphate transporter - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 05-Nov-1999
C:Accession: S68972; S49228
R:Helps, C.; Murer, H.; McGivan, J.
Eur. J. Biochem. 228, 927-930, 1995
A:Title: Cloning, sequence analysis and expression of the cDNA encoding a sodium-depende
A:Reference number: S68972; MUID:95255303; PMID:7737195
A:Accession: S68972
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-693 <HE2>
A:Cross-references: EMBL:X81699; NID:G547483; PIDN:CAA57345.1; PID:G547484

Query Match 79.3%; Score 2851; DB 2; Length 693;
Best Local Similarity 76.3%; Pred. No. 1.1e-193;
Matches 531; Conservative 82; Mismatches 73; Indels 10; Gaps 3;

QY	1	MAPWELGDAQNPDKYLEGAAQOQPTAPDKSKTNKNWTFAPVTKIELPSYSTATILID 60
DB	1	MAPWELGNSQPTSEKTVKADGEQSAKPEKAKETKDDTGTPTIKELVPSHSTATILIE 60
QY	61	EPTVEDDPMNLPTLOSGIKWSEKTKKILCFQGIKRLILLGLFYFVCSLDILSSA 120
DB	61	EPTVEDDPMNLPELKDUTGLKAWSEKTKKILCFQGIKRLILLGLFYFVCSLDILSSA 120
QY	121	FOLVGGKAGOFFPNSGINSNPLAGLVLGVLVTVLQSSSTSTSVWSVSSLLTVRAA 180
DB	121	FOLVGGKAVGKFFNNNSINSNPLAGLVLGVLVTVLQSSSTSTSVWSVASSLLPVHAA 180
QY	181	IPITMGANIGTSINTTIVALMQVDSFRAPAGATVHDFNNLSVLVLLPEVAVHYL 240
DB	181	IPITMGANIGTSINTTIVALMQVDSFRAPAGATVHDFNNLSVLVLLPEVAVHYL 240
QY	241	EITITOLIVESHFHNGEDAPDLIKVITKPFKTLIVOLDKKVISQIAMNDEKAKNSLVKI 300
DB	241	ERLTNLVESHFHNGEAPDLIKVITKPFKTLIQLDKSIINQIAMNDESVQNSMKIKI 300
QY	301	WCKTPTNTQINVTVPSTANTCSPSLCWTGDIQONWNTVYKENIAKQHFVNFHLPD 360
DB	301	WCKTPTNTVNTVTPSPENCTSPSLCWTGDIQONWNTVYKENIAKQHFVNFNLSD 360
QY	361	LAVGHTAILLSLVLCCLIMVILKLSVAGVATVTKKTINTDPPPPFPAWLCYLAAIL 420
DB	361	AVIGHTAILLSLVLCCLIMVILKLSVAGVATVTKKTINTDPPPPFPAWLCYLAAIL 420
QY	421	VGAGMTFTVQSSSVFTSALTPLIGITITIRAYPLTLGNSIGTITTTAILAALASPGNAL 480
DB	421	VGAGMTFTVQSSSVFTSALTPLIGITITIRAYPLTLGNSIGTITTTAILAALASPGSTL 480


```

Query Match      7.6%; Score 274.5; DB 2; Length 544;
Best Local Similarity 22.6%; Pred.No. 9.9e-12;
Matches 108; Conservative 52; Mismatches 115; Indels 203; Gaps 13;

Qy      94  FQIGIRL-ILLAGFLYFFVCKSLDILSSAFOLVGKGMAQGFSSNSIMSNPLLGLVGLVILV 152
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      11  FQIGGLGLFLGIKY-----MGDGLQMAAGDLRLDLKYT--TNPFGVLAGILV 60

Qy      153  TVLVQSSSTSIIVSVSWVSSLLVRAAIPITINGANIGTSITNTIVAMQVGRDSFERRA 212
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  TVLVQSSSTTVLTVGLVSAFGMTLQKAIGVINGANIGTTVTAFIIG----- 107

```

```

Db 108 -----IKUSEYSLPIIIVAGVALLFFPN----- 130
QY 273 LIVOLDKKVISQIAMNDEKANKSLKVIWCKTPTNKTQINTVTPSTANCTSPSLCWTDGI 332
Db 131 -----HKVN----- 135
QY 333 QNWTKMNVYKENIAKQHIIFVNHFLPDLAVGTILLISLVLCGLIMIVKILGSVLK- 391
Db 136 -----IGCVFFG-----GALFYGLIDLMGQMKP 159
QY 392 -----GOVATVIKKTINDPFPFPAWLTGYLAILVGCAGMTIVOSSVFTSALT 440
Db 160 LAGMESHETLQMS-----NPLGLLIGTIFTAVOSSATTGIIQ 202
QY 441 PLIGIGVITIERAYPLITGNSIGTTTTAILAASPINALRSSLQIALCHFNNISGILL 500
Db 203 ELYCGAGLDLQALPVLGSDNIGITTTAVLAIGASVAAKA-----AATHVIENLGAII 258
QY 501 WYPI--PFTRL-----PIRMAGLGNISAKY--RWFAPVYLIILFFFLIP 540
Db 259 FMLILPUFTSLFVAYLOGMFLNPMWITAVAHGPNITWTFIOFWIGAPAWLYTKLIP 316

```

RESULT 11
E95057.
Na/Pi cotransporter II-related protein SP0496 [imported] - Streptococcus pneumoniae (str: E95057).
C:Species: Streptococcus pneumoniae
C:date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:accession: E95057
R:Teitelin, H.; Nelson, K.E.; Paulsen, I.T.; Bisen, J.A.; Read, T.D.; Peterson, S.; Heiron, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; UID:21357209; PMID:11463916
A:Accession: E95057
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <R>
A:Cross-references: GB:AE005672; PID:gl4971968; GSFDB:GN00164; TIGR:SP:49800164
A:Experimental source: strain TIGR4

C:Genetics:
A:Gene: SP0496

Query Match 7.3%; Score 263.5; DB 2; Length 543;
Best Local Similarity 22.5%; Pred. No. 5.9e-11;
Matches 108; Conservative 53; Mismatches 134; Indels 185; Gaps 14;

```
QY 79 IKWSEDTGKILCFQGGIGRLILLGLFYFFVCSLDILSSAFQLVGKAGQFFNSIS 138
Db 3 INWQE-----ILFHF-----LGLGLFYISIKTNGDGLQAAGRLR--FYIDKY 45
QY 139 MSNPLGLGIVGLVTVLVSSTSTSIWMSVSSLLTVRAAIPITMGANIGTSINTIV 198
Db 46 TSNPFGLVGLGIMTALIQSSGVTVTVGLVSAGLLTRQAIGVGMGANIGTVTSPLI 105
QY 199 ALMQVDRSEFRRAFAGATVHDFNWLVLVLPVEVATHYLEITQLIVESHFKNGED 258
Db 106 G-----ILFHF-----LGLGLFYISIKTNGDGLQAAGRLR--FYIDKY 112
QY 259 APDLKLVITKPTKLIQVLDKKVISQIAMNDEKAKNSLVKWKCTFTNKTQINVTVPST 318
Db 113 ALPML-----NIGRL-----FGVGGIFPALNL--MSG 155
QY 319 ANCTSPSLCWTGDIQNTWTKNVTYKENIAKCOHIFVNFHPLDVLAVGTILLISLVLCGC 378
Db 134 -----NIGRL-----FGVGGIFPALNL--MSG 155
QY 379 LIMVKILGSLVKGQAVVIKKTINTDPPFPALWGTYLAILVAGMFTIVQSSSVFTSA 438
Db 156 MAPLKDL--QVFKDYMIELSKNPV-----LGVFVGTLTLIQSSATIGI 199
QY 439 LTPLIGIGVITIERAYPLTIGSNIGTITTTAILAALASFGNALRSSLOIALCHFFNISI 497
Db 200 LQNLVAGNLIDQGLPVLFGDNIGTITITAILASLANIAAKR---VAGAHVAFNIIGT 255
QY 498 -ILWVPIPT-----RLPIRMKGLGNISAKYRWFVYLIIFPF--LIP 540
Db 256 VFCVIFLVPFTVLHWFPEATNLAPENTIAFAHGTFTNTIVQPFPGALAYFVTKIIP 315
```

RESULT 12

G97926
conserved hypothetical protein spr0439 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: G97926
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21428245; PMID:11544234
A:Accession: G97926
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <R>
A:Cross-references: GB:AE007317; PIDN:AAK99243.1; PID:gl5458006; GSPDB:GN00174
C:Genetics:
A:Gene: spr0439

Query Match 7.3%; Score 262.5; DB 2; Length 543;
Best Local Similarity 22.7%; Pred. No. 7e-11;
Matches 109; Conservative 52; Mismatches 134; Indels 185; Gaps 14;

```
QY 79 IKWSEDTGKILCFQGGIGRLILLGLFYFFVCSLDILSSAFQLVGKAGQFFNSIS 138
Db 3 INWQE-----ILFHF-----LGLGLFYISIKTNGDGLQAAGRLR--FYIDKY 45
QY 139 MSNPLGLGIVGLVTVLVSSTSTSIWMSVSSLLTVRAAIPITMGANIGTSINTIV 198
Db 46 TSNPFGLVGLGIMTALIQSSGVTVTVGLVSAGLLTRQAIGVGMGANIGTVTSPLI 105
```

```
QY 199 ALMQVDRSEFRRAFAGATVHDFNWLVLVLPVEVATHYLEITQLIVESHFKNGED 258
Db 106 G-----ILFHF-----LGLGLFYISIKTNGDGLQAAGRLR--FYIDKY 112
QY 259 APDLKLVITKPTKLIQVLDKKVISQIAMNDEKAKNSLVKWKCTFTNKTQINVTVPST 318
Db 113 ALPML-----NIGRL-----FGVGGIFPALNL--MSG 155
QY 319 ANCTSPSLCWTGDIQNTWTKNVTYKENIAKCOHIFVNFHPLDVLAVGTILLISLVLCGC 378
Db 134 -----NIGRL-----FGVGGIFPALNL--MSG 155
QY 379 LIMVKILGSLVKGQAVVIKKTINTDPPFPALWGTYLAILVAGMFTIVQSSSVFTSA 438
Db 156 MAPLKDL--QVFKDYMIELSKNPV-----LGVFVGTLTLIQSSATIGI 199
QY 439 LTPLIGIGVITIERAYPLTIGSNIGTITTTAILAALASFGNALRSSLOIALCHFFNISI 498
Db 200 LQNLVAGNLIDQGLPVLFGDNIGTITITAILASLANIAAKR---VAGAHVAFNIIGT 255
QY 499 L--LWVPIPT-----RLPIRMKGLGNISAKYRWFVYLIIFPF--LIP 540
Db 256 VFCVIFLVPFTVLHWFPEATNLAPENTIAFAHGTFTNTIVQPFPGALAYFVTKIIP 315
```

RESULT 13

G83825
hypothetical protein BH1407 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83825
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, P.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83825
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BA05126.1; GSPDB:GN01
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1407

Query Match 6.9%; Score 247.5; DB 2; Length 543;

Best Local Similarity 22.1%; Pred. No. 8e-10;
Matches 104; Conservative 58; Mismatches 115; Indels 193; Gaps 15;

```
QY 90 ILCOFQIGRLIILGILFYFFVCSLDILSSAFQLVGKAGQFFNSISMSNPLGLGIVG 149
Db 9 LFWFQGLG--ILFGLIKY-----MGDGLQKVGRLRLDL--DKFTTFLMGVLAG 56
QY 150 VLVTVLVSSTSTSIWMSVSSLLTVRAAIPITMGANIGTSINTIVVALMOVGDRSEF 209
Db 57 IVTVVLOTSTGTTLTIGLVNAGFMTLKQAIGVGMGANIGTIVTAFIIG----- 106
QY 210 RRAFAGATVHDFNWLVLVLPVEVATHYLEITQLIVESHFKNGEDAPDLKLVITKP 269
Db 107 -----IKSEVALPIIANGAALIFFIKN----- 129
QY 270 FTKLIQVLDKKVISQIAMNDEKAKNSLVKWKCTFTNKTQINVTVPSTANCTSPSLCWT 329
Db 130 -----KKV----- 132
QY 330 DGIQNTWTKNVTYKENIAKCOHIFVNFHPLDVLAVGTILLISLVLCGLIMVILGVS 389
Db 133 -----NNIG--QVIF-----GFTLFYGLNTM----- 152
QY 390 LKGOVATVTKTINTDPPFPALM-----CYLAILVAGMFTIVQSSSVFTSALTPLI 443
Db 153 --GEGNLPUREL-----QAPADLVSMENPILGLIGTITFAVQSSSASTGLGQQLY 204
QY 444 GIGVITIERAYPLTIGSNIGTITTTAILAALASFGNALRSSLOIALCHFFNISG--ILLW 501
```

Db 205 DQGMDFALPVLFGNIGTITAVLAAGASVAARA-----ALTHVFNLGTTIIVLI 260
QY 502 YEIPPTR-----LPIRMAGLGNIS--AKYRWFAPYLLI 534
Db 261 IIIPTHTIAXLAUEVPALEARNPTIAPAGIFNVNTIIQPFPGIILAIIV 310

RESULT 14
E90554
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C;Accession: E90554
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: E90554
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-569 <KUR>
A;Cross-references: GB:AL445566; PID:g14089755; PIDN:CAC13514.1; GSPDB:GN00153
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU 3410
A;Genetic code: SGC3

Query Match 6.7%; Score 242.5; DB 2; Length 569;
Best Local Similarity 20.7%; Pred. No. 1.9e-09;
Matches 109; Conservative 84; Mismatches 122; Indels 211; Gaps 23;

QY 91 LCFQOGIRLILIGFLYFVCSLDILGSAPQLVGGKMGQFPSSNSIMSNPLGLVIGV 150
Db 18 LGWEGI--LLVFG-ISVPLMSIKMSTSRNVGSEKFKIL--LSRSPRIIGIAGP 72

QY 151 LVTVLQSSSTSTSVVMSVSSLTVRAAIPIMGANIGTSITNTVIMVQGVDRSEFR 210
Db 73 AFTSMIQSDGAVALLGALAFIDLGATFAGVIGANVTATTSIIVALEQ----- 124

QY 211 RAFAGATVHDFNWLVLVLLPEVATHYLEITQLIVSFHFKNGEDAPDLLKVTYPF 270
Db 125 -----YFKISNYLLM-----CVIGGFIF----- 143

QY 271 TKLIVQLDKKVISQIAMDDEKAKNSLVKIWKCTFTNKTQINVTVPSTANCTSPSLCWD 330
Db 144 -----LFFK----- 148

QY 331 GIONWTKNVTYKENIAKQHIFVNFHPLDLAVGTILLILSLVLCGLIMIVKILGSLV 390
Db 149 --EKWT-----KIG-FLIFSI-----GMIFLGLKILGSV 175

QY 391 KGOVATVTKTINTDPPPPA-----WLTGYLAILVAGMTFIVQSSSVTSALTPLI 443
Db 176 K-----VLTSQVPSDFKFGQDSASNNISFFPSI-----ALTALFQSSATIAIYQAIL 227

QY 444 --GIGVITIRAYPLTIGSNIGTTTTAILAALSPGNALRSSLOIALCHFFPNISGILM 501
Db 228 SGENILSLSGALVLCANIGTTITAILIAFVS-GB--KNSRIAISWLFNSIALIA 284

QY 502 YPI-----PFTPLIRMAKGLNISAKYRW-----PAVFYLIIFPFIPIPTVFGLS 547
Db 285 NSLVGVAAPFVRL---IVGENEIAA---WQLSIAHLLENIFLAVVFFFIKQLV---- 334

QY 548 LAGWELVGVGVFVFIILVLCRLQRCPR-----VLPKQLQN 588
Db 335 ---W-----LCFVLKKSEDEKTEGMPLPKLVN 360

RESULT 15

H71283

conserved hypothetical integral membrane protein TP0771 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C;Accession: H71283
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McDc
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: H71283
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-593 <COL>
A;Cross-references: GB:AE001248; GB:AE000520; NID:g3323074; PIDN:AAC65739.1; PID:g332307
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0771

Query Match 6.1%; Score 218.5; DB 2; Length 593;
Best Local Similarity 20.7%; Pred. No. 9.9e-08;
Matches 111; Conservative 78; Mismatches 149; Indels 199; Gaps 20;

QY 31 KSKETKNNT--EAPVTKIELLPSYSTATLIDPTEVDVDPNWLPTLODSGKMSERDTK 87
Db 9 KERUTNKHTSVRSVPLRTIDTYPLFL-----W-----WAPMELV 44

QY 88 GKILCFQOGIRLILIGFLYFVCSLDILGSAPQLVGGKMGQFPSSNSIMSNPLGLV 147
Db 45 AVLL-----QGIGS-----LGFMLP--GMKLSGDGQKGTGESLHKIL--QVMTSNCFFAVL 92

QY 148 TGVLTVLVQSSSTSTSVVMSVSSLTVRAAIPIMGANIGTSITNTVIMVQGVDRS 207
Db 93 TGMATVAIVQSSGATVTMTISFINAGVLSITQSIGVILGANIGTTVTAMIVAV----- 146

QY 208 EERRAFAGATVHDFNWLVLVLLPEVATHYLEITQLIVSFHFKNGEDAPDLLKVTIT 267
Db 147 -----GFOFK----- 151

QY 268 KFTKLIIVQLDKKVISQIAMDDEKAKNSLVKIWKCTFTNKTQINVTVPSTANCTSPSLC 327
Db 152 -----LASTAVPAP----- 160

QY 328 WTDGIONWTKNVTYKENIAKQHIFVNFHPLDLAVGTILLILSLVLCGLIMIVKILG 387
Db 161 ---GIGYF---LTFPKHLHK-----CHLGEISMGFGLFTGL---GLLSSLIPLLS 202

QY 388 ----SVLKGQVATVTKTINTDPPPPAWLTGYLAILVAGMTFIVQSSSVTSALTPLI 443
Db 203 VERLSFLKIAVED-----RALSVFVGLLSGFVLTVILHSSSAFTAIVLTM 248

QY 444 GIGVITIRAYPLTIGSNIGTTTTAILAALSPGNALRSSLOIALCHFFPNISG----IL 499
Db 249 FGGVIGVEFAASVLSNVGVTIDAAIAIGSKLNARRA-----AAHVLFNVFGALVFLM 304

QY 500 LWPYIPFTRPIRMAK--GLGNISAKYRW--AVFYLIIFPFIPIPTVFGLSAGWRV 553
Db 305 FPHPV-LALLCVLTTPKNSGFDNITVRLALFHSMENIVNTIIVFPFTKHLAAVFEWLI 360

Search completed: August 26, 2003, 15:53:41
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2003, 15:41:38 ; Search time 26 Seconds
(without alignments)
1246.208 Million cell updates/sec

Title: US-10-052-664-1

Perfect score: 3597

Sequence: 1 MAPWPELGDAQPNPKYLEG.....SREAQGEVPASDSKTECTAL 689

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127963 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127963

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1721	47.8	637	1 NPT2 RAT	Q06496 rattus norv
2	1717	47.7	639	1 NPT2 SHEEP	Q97704 oviss aries
3	1698.5	47.2	639	1 NPT2 HUMAN	Q06495 homo sapien
4	1681.5	46.7	642	1 NPT2 RABIT	Q28620 o renal sod
5	1667	46.3	637	1 NPT2 MOUSE	Q08025 mus musculu
6	203	5.6	310	1 Q0EW BACSU	P54463 bacillus su
7	165	4.6	543	1 XJBB ECOLI	P32683 escherichia
8	144.5	4.0	543	1 XJBB SALT	P40730 salmonella
9	139.5	3.9	630	1 YD13 SCHPO	Q92341 schizosacch
10	129	3.6	654	1 PSTA MYCGE	P47651 mycoplasma
11	128.5	3.6	556	1 OCN3 HUMAN	O75751 homo sapien
12	128	3.6	512	1 COX1 LOCMI	Q36421 locusta mig
13	127.5	3.5	524	1 LCTP BACHD	Q9k529 bacillus ba
14	127	3.5	551	1 OCN3 MOUSE	Q9wtw5 mus musculu
15	125	3.5	547	1 SNG1 YEAST	P46950 saccharomyc
16	123.5	3.4	551	1 OCN3 RAT	Q88446 rattus norv
17	121.5	3.4	896	1 BOSS DROME	P22815 drosophila
18	121	3.4	508	1 COX1 BLAGE	Q36724 blattella g
19	118	3.3	464	1 NORM PASMU	Q9cmz9 pasteurella
20	116	3.2	448	1 NMTH LISMO	Q8y773 listeria mo
21	116	3.2	1021	1 S123 HUMAN	P55017 homo sapien
22	115	3.2	651	1 PSTA MYCPN	P75185 mycoplasma
23	114	3.2	512	1 COX1 RHISA	Q99818 rhinipcephal
24	114	3.2	517	1 COX1 MYXGL	Q21079 myxine glut
25	114	3.2	893	1 BOSS DROVI	Q24738 drosophila
26	113	3.1	448	1 NMTH LISIN	Q92bt1 listeria in
27	112.5	3.1	518	1 COX1 SQAC	Q92z52 squalus aca
28	111.5	3.1	517	1 COX1 SCYCA	O79403 scyllorhinu
29	110.5	3.1	613	1 NUSM POLOR	Q95918 polypetereus
30	110	3.1	429	1 URAA ECOLI	P33780 escherichia
31	110	3.1	475	1 SECY MYCGE	P47416 mycoplasma
32	110	3.1	514	1 COX1 ANOGA	P34838 anopheles g
33	110	3.1	812	1 MPD_EUCAI	P57381 buchmera ap

34	110	3.1	1132	1 NUT1 YEAST	P53114 saccharomyc
35	109.5	3.0	388	1 YUBA_BACSU	Q32086 bacillus su
36	109	3.0	514	1 COX1 ANOQU	P33504 anopheles q
37	109	3.0	543	1 HTI_TRIVI	Q27115 trypanosoma
38	109	3.0	607	1 NUSM MOUSE	P03921 mus musculu
39	108.5	3.0	369	1 Y316 MYCPN	Q50361 mycoplasma
40	108.5	3.0	656	1 SL54 MOUSE	Q9et37 mus musculu
41	108	3.0	356	1 CKF8 MACMU	Q97665 macaca mula
42	107.5	3.0	687	1 AFUB_ACTPL	Q44123 actinobacil
43	107.5	3.0	739	1 S6A5 RAT	P58295 rattus norv
44	107.5	3.0	1476	1 CFTR MOUSE	P26361 mus musculu
45	106.5	3.0	197	1 CSP_XENLA	O42196 xenopus lae

ALIGNMENTS

RESULT 1	
NPT2 RAT	
ID_NPT2_RAT	STANDARD; PRT; 637 AA.
AC	Q06456;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-FEB-1995 (Rel. 31, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Renal sodium-dependent phosphate transport protein 2 (Sodium/phosphate cotransporter 2) (Na(+)/Pi cotransporter 2) (Renal sodium-phosphate cotransporter 2) (Renal Na(+)-dependent phosphate cotransporter 2).
DE	SLC34A1 OR SLC17A2.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Kidney;
RX	MEDLINE=93317607; PubMed=8327470;
RA	Magagnin S., Werner A., Markovich D., Sorribas V., Stange G., Biber J., Murer H.;
RA	"Expression cloning of human and rat renal cortex Na/Pi cotransport.";
RL	Proc. Natl. Acad. Sci. U.S.A. 90:5979-5983(1993).
RN	[2]
RP	DISULFIDE BONDS, AND REVISED TOPOLOGY.
RX	MEDLINE=20385426; PubMed=10926678;
RA	Lambert G., Forster I.C., Biber J., Murer H.;
RT	"Cysteine residues and the structure of the rat renal proximal tubular type II sodium phosphate cotransporter (rat Napi IIA).";
RL	J. Membr. Biol. 176:133-141(2000).
CC	-!- FUNCTION: MAY BE INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE INTO CELLS VIA NA+ COTRANSPORT IN THE RENAL BRUSH BORDER MEMBRANE.
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.
CC	-!- TISSUE SPECIFICITY: Kidney.
CC	-!- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).
CC	EMBL; L13257; AAC37608.1; -.
DR	PIR; A48189; A48189.
DR	InterPro; IPR003841; Na/Pi_cotranspt.
DR	Fram; PF02690; Na_Pi_cotrans; 2.
DR	TIGRfam; TIGR01013; 2a58; 1.
KW	Transport; Symport; Sodium transport; Transmembrane; Glycoprotein; Phosphorylation.
FT	DOMAIN 1 103 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 104 125 MI (POTENTIAL).
FT	DOMAIN 126 145 EXTRACELLULAR (POTENTIAL).

QY 222 FNLVSLVLLPVEVATHYLEITOLIVESHFKNGEDAPLLKVTIKPTKLIIVOLDKV 281
 DB 226 FNLVSLVLLPVEVATHYLEITOLIVESHFKNGEDAPLLKVTIKPTKLIIVOLDKV 285
 QY 282 ISQIANDDEKANKSLVVKWCKTFTNTQKINVTVPSTANTCTSPSLCWTGDIQNTWKV 341
 DB 286 ITSLSAGDESRLSHLRVWC--YNPTEVPTMPRAEANTSRML-----RNAT 332
 QY 342 YKENTAKOHLVFNPHLPDLAVCTILLILSLVLCGLIMVLKLSVKGQAVIKKT 401
 DB 333 ----LEKCNHIFVDTGLPDLAVGLILLAGSLALCTCLILLVXNLSLKGQAVIKQV 388
 QY 402 INTDPPFFAMLTGLYAILVAGMTFIVQSSVFTSALTPLIGVITITERAYPLTIGSN 461
 DB 389 INTDPPFTWATGYFAMVVGASMTFVQSSVFTSALTPLIGVITITERAYPLTIGSN 448
 QY 462 IGTITTAIALAGNALRSSLOIALCHFFNFISGILLWYPIPTLPIRMALGNIS 521
 DB 449 IGTITTAIALAGPREKLSSAFQIALCHFFNFISGILLWYPIPTLPIRMALGNIS 508
 QY 522 AKYRFAYVYLLIIFEFILPTVGLSLAGRVLCVGVVFFIILVLCRLQSRCPV 581
 DB 509 AKYRFAYVYLLIIFEFILPTVGLSLAGRVLCVGVVFFIILVLCRLQSRCPV 568
 QY 582 LPKLQNNFLEPMRSIKPDVAVSKFTGCFQMRCCOCC 622
 DB 569 LPKLQNNFLEPMRSIKPDVAVSKFTGCFQMRCCOCC 602
 RESULT 3
 NPT2 HUMAN STANDARD; PRT; 639 AA.
 ID NPT2 HUMAN STANDARD; PRT; 639 AA.
 AC Q06455;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Renal sodium-dependent phosphate transporter protein 2 (Sodium/phosphate cotransporter 2) (Na+/Pi cotransporter 2) (Renal sodium-phosphate transport protein 2) (Renal Na(+)-dependent phosphate cotransporter 2).
 GN SL34A1 OR SL317A2 OR NPT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 OC NCBI_TaxID=9606;
 EN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=93317607; PubMed=8327470;
 RA Magagnoli S., Werner A., Markovich D., Sorribas V., Stange G., Biber J., Murer H.;
 RA "Expression cloning of human and rat renal cortex Na/Pi cotransport."; Proc. Natl. Acad. Sci. U.S.A. 90:5979-5983 (1993).
 RL "Expression cloning of human and rat renal cortex Na/Pi cotransport."; Proc. Natl. Acad. Sci. U.S.A. 90:5979-5983 (1993).
 CC -!- FUNCTION: MAY BE INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE INTO CELLS VIA NA+ COTRANSPORT IN THE RENAL BRUSH BORDER MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: KIDNEY, AND LUNG.
 CC -!- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L13258; AAA36354.1; --
 DR PIR: B48189; B48189.
 DR Genbank: U01019; SL34A1.
 DR MIM: 182309; --

DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0015321; P: sodium-dependent phosphate transporter acti. . .; TAS.
 DR GO; GO:0007589; P: fluid secretion; TAS.
 DR GO; GO:0006796; P: phosphate metabolism; TAS.
 DR GO; GO:0006817; P: phosphate transport; TAS.
 DR InterPro; IPR003841; Na/Pi cotranspt.
 DR Pfam; PF02690; Na Pi cotrans; 2.
 DR TIGRFAMs; TIGR01013; 2a58; 1.
 KW Transport; Symport; Sodium transport; Transmembrane; Glycoprotein; Phosphorylation.
 KW DOMAIN 1 103 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 104 125 M1 (POTENTIAL).
 FT DOMAIN 126 145 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 146 163 M2 (POTENTIAL).
 FT DOMAIN 164 165 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 166 185 M3 (POTENTIAL).
 FT DOMAIN 186 347 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 348 370 M4 (POTENTIAL).
 FT DOMAIN 371 412 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 413 436 M5 (POTENTIAL).
 FT DOMAIN 437 466 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 467 487 M6 (POTENTIAL).
 FT DOMAIN 488 513 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 514 534 M7 (POTENTIAL).
 FT DOMAIN 535 539 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 540 561 M8 (POTENTIAL).
 FT DOMAIN 562 639 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 639 669 LEUCINE-ZIPPER.
 FT MOD RES 508 508 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 639 AA; 68937 MW; 65D21D968C35D61B CRC64;
 Query Match 47.2%; Score 1698.5; DB 1; Length 639;
 Best Local Similarity 58.7%; Pred. No. 3.4e-108;
 Matches 348; Conservative 82; Mismatches 118; Indels 45; Gaps 9;
 QY 43 PVTKEILLPSVSTATID-----EPTVDPPWNL-----TLQDSGKSKSERDTGKILCF 93
 DB 42 PGTSAFAPSLGVALAHTCPGVLERHEP--LPKALAE-----EQPSRLVPK 93
 QY 94 FQIGRLI-----LLGLFYFVCSLDLSSAFQVGGKAGOFFNSSTMSNPLGLVIG 149
 DB 94 LEQAGAMLLVPLMLFYLFCVSLDMLSSAFQAGKAGDIFKNALSNPVLAVVG 153
 QY 150 VLVTVLVQSSSTSTSVSWSSSLLTVRAIPIINGANIGTSITNTVALMQVDSRSEF 209
 DB 154 ILVTVLVQSSSTSTSVSWSSSLLTVRAIPIINGANIGTSITNTVALMQVDSRSEF 213
 QY 210 REAPAGATVDEDFNLSVLVLPVEVATHYLEITOLIVESHFKNGEDAPLLKVTIKP 269
 DB 214 REAPAGATVDEDFNLSVLVLPVEVATHYLEITOLIVESHFKNGEDAPLLKVTIKP 273
 QY 270 FTKLIVOLDKKVISQIANDDEKANKSLVVKWCKTFTNTQKINVTVPSTANTCTSPSLCWT 329
 DB 274 FTKLIVOLDKKVISQIANDDEKANKSLVVKWCKTFTNTQKINVTVPSTANTCTSPSLCWT 320
 QY 330 DGIQNTWKVTKNTKNIACQHI FVNPHLPDLAVCTILLILSLVLCGLIMVLKLSVKGQAVIKKT 389
 DB 321 EANGSQTIGNAT-----MEKCNHIFVDTGLPDLAVGLILLAGSLALCTCLILLVXNLSL 376
 QY 390 LKGQAVTKTKTINTDPPFFAMLTGLYAILVAGMTFIVQSSVFTSALTPLIGVIT 449
 DB 377 LKGQAVTKTKTINTDPPFFAMLTGLYAILVAGMTFIVQSSVFTSALTPLIGVIT 436
 QY 450 TERAYPLTIGSNIGTITTAIALAGNALRSSLOIALCHFFNFISGILLWYPIPTLPIRMALGNIS 509
 DB 437 TERAYPLTIGSNIGTITTAIALAGNALRSSLOIALCHFFNFISGILLWYPIPTLPIRMALGNIS 496
 QY 510 PIRMAKGLGNISAKYRFAYVYLLIIFEFILPTVGLSLAGRVLCVGVVFFIILVLCRLQSRCPV 569
 DB 497 PIRMAKGLGNISAKYRFAYVYLLIIFEFILPTVGLSLAGRVLCVGVVFFIILVLCRLQSRCPV 556

QY 570 CLELQSCPRVLPKQLWNNELPLNRSILKPDVAVSKFTGCFQMRCCOCC 622
 DB 557 LINVLOSREPGLPKWLQTDWFLRWNHSLKPLDLHLITRAT-----ICCAR 602

RESULT 4
 NPT2 RABIT
 ID_NPT2 RABIT STANDARD; PRT; 642 AA.
 AC Q28620;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Renal sodium-dependent phosphate transport protein 2 (Sodium/phosphate cotransporter 2) (Na⁺/Pi cotransporter 2) (Renal sodium-phosphate transport protein 2) (Renal Na⁺)-dependent phosphate cotransporter 2) (Napi-6).
 DE (Napi-6).
 GN SLC34A1 OR SLC17A2.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95251013; PubMed=7733319;
 RA Verri T., Markovich D., Perego C., Norbis F., Stange G., Sorribas V., Biber J., Murer H.;
 RA "Cloning of a rabbit renal Na-Pi cotransporter, which is regulated by dietary phosphate.";
 RT J. Biol. Chem. 268:6226-6233 (1993).
 RL Am. J. Physiol. 268:F626-F633 (1995).
 CC -!- FUNCTION: MAY BE INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE INTO CELLS VIA NA+ COTRANSPORT IN THE RENAL BRUSH BORDER MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; U20793; AAA77682.1; --
 DR PIR; I46534; I46534.
 DR InterPro; IPR003841; Na/Pi_cotransp.
 DR Pfam; PF02690; Na_Pi_cotrans; 2.
 DR TIGRFAMs; TIGR01013; 2a58; 1.
 KW Transport; Symport; Sodium transport; Transmembrane; Glycoprotein; Phosphorylation.
 KW DOMAIN 1 106 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 107 128 M1 (POTENTIAL).
 FT DOMAIN 129 146 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 149 166 M2 (POTENTIAL).
 FT DOMAIN 167 188 M3 (POTENTIAL).
 FT TRANSMEM 169 188 M3 (POTENTIAL).
 FT DOMAIN 189 350 M4 (POTENTIAL).
 FT TRANSMEM 351 373 M4 (POTENTIAL).
 FT DOMAIN 374 415 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 416 439 M5 (POTENTIAL).
 FT DOMAIN 440 469 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 470 490 M6 (POTENTIAL).
 FT DOMAIN 491 516 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 517 537 M7 (POTENTIAL).
 FT DOMAIN 538 542 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 543 564 M8 (POTENTIAL).
 FT DOMAIN 565 642 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 642 655 LEUCINE-ZIPPER.
 FT DOMAIN 655 372 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT MOD_RES 511 511 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 333 333 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 642 AA; 69431 MW; A4885C001642A8B CRC64;
 Query Match 46.7%; Score 1681.5; DB 1; Length 642;
 Best local similarity 58.6%; Pred. No. 4.9e-107;
 Matches 349; Conservative 77; Mismatches 119; Indels 51; Gaps 10;
 QY 43 PVTKEILPSSYSTATLD-----RPTVEDDPWMLPT-----LQDSGIKWS--RDT 86
 DB 45 PGTSAFAPSLSPVALTERGCPYGEARHEP--LPKALAEEOQPESG--WAQELRPT 100
 QY 87 KKKILCFPGIGRILLLGLFLYFVCSLDILSSAFQLVGKMGAGOFFSNSSIMNPLGL 146
 DB 101 AMTLL-----KLPLMVTFLYFVCSLDVLSAFQLVGKMGAGOFFSNSSIMNPLGL 153
 QY 147 VIGVLTAVLVQSSSTSTSIIVSNVSSLLTYRAAIPILMANGNIGTSITNTIIVLMQVDR 206
 DB 154 VVGILVTVLVQSSSTSTSIIVSNVSSLLTYRAAIPILMANGNIGTSITNTIIVLMQVDR 213
 QY 207 SEFRRAFAGATVHDFFNWLVLVLPVAVATHYLEITVLIIVSFPHKNGEDAPLLKVI 266
 DB 214 TDFRRAFAGATVHDFFNWLVLVLPVAVATHYLEITVLIIVSFPHKNGEDAPLLKVI 273
 QY 267 TKPFTKLIIVOLDKKVISQIAMNDEKAKNSLVKWKCTFTNKTQINVTVPSTANCTPSL 326
 DB 274 TEPFTKLIIVOLDKKVISQIAMNDEKAKNSLVKWKCTFTNKTQINVTVPSTANCTPSL 327
 QY 327 CWTGDIQNTWNTYKENIAKOHFVNFLPDVAVTLLILSLVLCGLIMVKIL 386
 DB 328 SRTHG--NATVE-----KCNHIFVDVQPLDVLAVGLLGLAGSLVLLCTCHILVKNL 376
 QY 387 GSVLKGQVATVKKTIINTDPPFPFAMVLCYLAIVAGAMTFIVQSSSVFTSALTPLIG 446
 DB 377 NSLLKGQVATVKKTIINTDPPFPFAMVLCYLAIVAGAMTFIVQSSSVFTSALTPLIG 436
 QY 447 VTITRAVPLTGNIGTTTTTALALASPGRALRSSLOALCHFFPNTSGILLWYPIPF 506
 DB 437 VISIERAVPLTGNIGTTTTTALALASPGRALRSSLOALCHFFPNTSGILLWYPIPF 496
 QY 507 TRLPRMAGKGNISAKYRFAVYVLIFFFLIPLTVRGLSLAGWLVGVGVVVFII 566
 DB 497 TRLPRMAGKGNISAKYRFAVYVLIFFFLIPLTVRGLSLAGWLVGVGVVVFII 556
 QY 567 LVLCRLQLQSCPRVLPKQLWNNELPLNRSILKPDVAVSKFTGCFQMRCCOCC 622
 DB 557 FVVLINVLSQSPGLEPKWLQTDWFLRWNHSLKPLDLHLITRAT-----LCCSR 605

RESULT 5
 NPT2 MOUSE
 ID_NPT2 MOUSE STANDARD; PRT; 637 AA.
 AC Q60825; Q62110; Q62111; Q62112; Q62113; Q62114; Q62115; Q62116;
 AC Q62564;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Renal sodium-dependent phosphate transport protein 2 (Sodium/phosphate cotransporter 2) (Na⁺/Pi cotransporter 2) (Renal sodium-phosphate transport protein 2) (Renal Na⁺)-dependent phosphate cotransporter 2) (Napi-7).
 DE (Napi-7).
 GN SLC34A1 OR SLC17A2 OR NPT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/c; TISSUE=Kidney cortex;
 RX MEDLINE=96006598; PubMed=7478940;
 RA Hartmann C.M., Wagner C.A., Busch A.E., Markovich D., Biber J., Laug F., Murer H.;
 RA "Transport characteristics of a murine renal Na/Pi-cotransporter.";
 RL Pflugers Arch. 430:830-836 (1995).

[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney cortex;
RA MEDLINE=94350180; PubMed=8070635;
RX Collins J.F., Ghishan P.K.;
RT "Molecular cloning, functional expression, and in
RT situ hybridization of the renal sodium phosphate (Na+/Pi)
RT transporter in the control and hypophosphatemic mouse.";
RL FASEB J. 8:862-869 (1994).
[3]
RP SEQUENCE OF 1-94; 110-177; 179-242; 254-312; 313-318; 332-333; 335-337
RP AND 349-485 FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=96293539; PubMed=8693007;
RA Hartmann C.M., Hewson A.S., Kos C.H., Hilfiker H., Soumounou Y.,
RT Murer H., Tenenhouse H.S.;
RT "Structure of murine and human renal type II Na+-phosphate
RT cotransporter genes (Npt2 and NPT2).";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7409-7414 (1996).
CC -!- FUNCTION: MAY BE INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE INTO
CC CELLS VIA NA+ COTRANSPORT IN THE RENAL BRUSH BORDER MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Kidney.
CC -!- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L33878; AAC42026.1; -;
DR EMBL; U56664; AAC52684.1; ALT_INIT.
DR EMBL; U56665; AAC52685.1; -;
DR EMBL; U56666; AAC52686.1; -;
DR EMBL; U56667; AAC52687.1; -;
DR EMBL; U22465; AAC52361.1; -;
DR EMBL; U56668; AAC52688.1; -;
DR EMBL; U56669; -; NOT_ANNOTATED_CDS.
DR EMBL; U56670; AAC52689.1; -;
DR EMBL; U56671; AAC52690.1; -;
DR EMBL; U56672; AAC52691.1; -;
DR EMBL; U56673; AAC52692.1; -;
DR MGD; MGI:1345284; SLC34a1.
DR InterPro; IPR003841; Na/Pi_cotransp.
DR Efcg; PF02690; Na_Pi_cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
KW Transport; Symport; Sodium transport; Transmembrane; Glycoprotein;
KW Phosphorylation.
KW DOMAIN 1 103 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 104 125 M1 (POTENTIAL).
FT DOMAIN 126 145 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 146 163 M2 (POTENTIAL).
FT DOMAIN 164 165 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 166 185 M3 (POTENTIAL).
FT DOMAIN 186 345 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 346 368 M4 (POTENTIAL).
FT DOMAIN 369 410 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 411 434 M5 (POTENTIAL).
FT DOMAIN 435 464 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 465 485 M6 (POTENTIAL).
FT DOMAIN 486 511 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 512 532 M7 (POTENTIAL).
FT DOMAIN 533 537 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 538 559 M8 (POTENTIAL).
FT DOMAIN 560 637 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 637 677 LEUCINE-ZIPPER.
FT FT MOD_RES 677 696 PHOSPHORYLATION (BY PKC) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. .) (POTENTIAL).
FT FT CONFLICT 324 324 G -> A (IN REF. 2).

FT	CONFLICT	345	A -> D (IN REF. 2).
FT	CONFLICT	390	MSSRSSSTQT -> NIVQKVINTD (IN REF. 2).
FT	CONFLICT	444	G -> D (IN REF. 2).
SQ	SEQUENCE	637 AA;	1796935971231C09 C064;

Query Match 46.3%; Score 1667; DB 1; Length 637;
Best Local Similarity 58.4%; Pred. No. 4,7e-106;
Matches 340; Conservative 80; Mismatches 126; Indels 36; Gaps 5

Qy	43	PVTKIELLPYSVTALID-----RPEVDVDPWNLT-LQDSGIKWSEDTGKILCFQ 95
Dd	42	FETSYAISLSPVUTEHSQPCGVLECHD--LPTKLAQEERQPRLSQKLAVGT 99
Qy	96	GIGRLILLGLFFYCSHDILSAPQLVGCKWAGOFFSNSIMNPNLLGLVIGLVTVTL 155
Dd	100	KLKVPMLAFLFYCLSDLVSNFQLAGKVAGDIFKNAILSPNVAGVLGVITVL 159
Qy	156	VQSSTSTSIVMSVSSSLTVRAAPIIMGANIGTSINTTIVALMQVDRSEFRPAPAG 215
Dd	160	VQSSTSTSIVMSVSGILLEVSAPIPMGSMIGTSVNTTIVALMQAGRTDFRRAFAG 219
Qy	216	ATVHDFNMLSVLALLPEVAATHYLEIITOLIVESHFKNGEDADIALKVIITEPTKLIV 275
Dd	220	ATVHDFNMLSVLULLPERATGYLHVHTGLVASHFNIRGERDAPDLLKVIITEPTKLII 279
Qy	276	QLDKVISIQAMNDEKAKNSKIWIWCKTFNKTNQINV---VPSTANCTPSLCWTDG 331
Dd	280	QLDKSVITSANGDESRLNHSILIHWCPDPTTEASTMSRVEAIGSLANT----- 330
Qy	332	TONTMKNVYKENTAKOCHFVNPHPLDIAGVTILLISLVLGCGLIMTVIKLGSLVK 391
Dd	331	-----MEKCNHFVDYGLPALAVGLILLAGSLVWLCTCLILLAKMNSLLK 376
Qy	392	GOVATVIKTINDPPFPAMIYGVAILVCGMTTWIVOSSVPTSALTPLIGHGVITIE 451
Dd	377	GOVAMSSRSSTQTPFPAPTWTGTGFAMVVGASMTFWOSSVFTSAITPLIGHGVISIE 436
Qy	452	RAYPLTGSMIGTTTVAIALAASPINALRSSIQIALCHFFNFNISGILLWYPFTLRPI 511
Dd	437	RAYPLTGSMIGTTTVAIALAASPREKLSSFOIALCHFFNFNISGILLWYPLCTLPI 496
Qy	512	RMAKGLNISKARYNPFAVYLIIFFLIPIUPTVGLSLAGRWLVGVGPVFIIILVLCL 571
Dd	497	RMAKALKRKTRYKRFAYVLIIIVCFLLIPSILVFGISMAGWQMVGVPFCGALLAFVVLV 556
Qy	572	RILOSRCFVLPKLQWNFIPLWRSIKPWDVAWSKFTGCF 613
Dd	557	NVLORSFGHLPKWLQTDWDLFRWMHSLQDGLITRALICY 598

RESULT 6

QYEW_BACSU	STANDARD;	PRT;	310 AA.
ID	YOEW_BACSU		
AC	P54463;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DE	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Hypothetical protein yqew.		
GN	YOEW.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168 / JH642;		
RX	MEDLINE=97124195; PubMed=8965908;		
RA	Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M., Kobayashi Y.;		
RA	"Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes.";		
RL	Microbiology 142:3103-3111(1996).		
RN	[2]		


```

CC -!- SIMILARITY: TO B.SUBTILIS YQEW.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00006; AAC43114.1; -.
DR EMBL; AE000475; AAC76990.1; -.
DR EMBL; AE015417; AAN45511.1; -.
DR PIR; C65209; C65209.
DR EcoGene; EG11919; yjbb.
DR InterPro; IPR003841; Na/Pi cotranspt.
DR InterPro; IPR004633; Napi cotransptII.
DR Pfam; PF02690; NaPi_cotrans; 1.
DR TIGRFAMs; TIGR01013; 2a58; 1.
DR TIGRFAMs; TIGR00704; Napi cotrn rel; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
SQ SEQUENCE 543 AA; 59466 MW; BA97149CC41A093F CRC64;

Query Match 4.6%; Score 165; DB 1; Length 543;
Best Local Similarity 20.6%; Pred. No. 0.00058;
Matches 94; Conservative 58; Mismatches 134; Indels 170; Gaps 14;

Qy 100 LILLIGFLYFFVCSLDLSAFLQVGGKAGQFSSNSINPNLLGLVIGLVTVLVQSS 159
Db 4 LHLLSAVALLVWGTHTVGTGMVFGARLTVLSR-SVEKTP-LAFCAGIGTALVQSS 61
Qy 160 STSTSVVSMVSSSLTVAAPITWGANIGTSTNTIVAMQVGRSEFRAPAGATVH 219
Db 62 NATTLVTSFVAQDLVALPALVILGADVGTALMARILTF----- 102
Qy 220 DFFNWLVLVLPVEVATHYLEITQLIVESHFKNGEDAPDLLKVTYKPTKLIQVLDK 279
Db 103 -DLSWLSPLLI-----FTGVIFPLGR 122
Qy 280 KVISQIANDKAKKSLVKWCKTFNTKTQINTVVPSTANCTSPSLQWTDGIQNTWTKN 339
Db 123 K-----QSRAGOLGRV----- 133
Qy 340 VTYKENIAKQHIFVNFPLDLAGVTILLILSLVLCGLIMIVKILGSLVKGVATVIK 399
Db 134 -----GIGLILLALEL-----IVQAVTPTQANGVQVI- 163
Qy 400 KTIINDPFPFPAWLTG--YLAILVAGMTFTVQSSVFTSALTPPLIGIGVITIERAYPLT 457
Db 164 -----FASLTGILLDALIGAMFAIISYSLAAVLLTATLTAAGIISFPVALCLV 213
Qy 458 LGSNIGTITTAALASPCNARLSLQIALCHFFNIGLILWYIPTRIPTRIMAKGL 517
Db 214 IGANIGSGLLAMLNNSAANAAR-----VALGSLLEFKLVGSLI--ILPFVHL---LAETM 264
Qy 518 GNISAKYR-----WFAVEY-----LIIFPLIPLTVF 544
Db 265 GKLSLPKALVIYFHFVYNVRLVLMPLFVDPMARF 300

RESULT 8
YJBB SALTY
ID YJBB SALTY STANDARD; PKT; 543 AA.
AC P40730;
DT 01-FEB-1995 (Rel. 31, Created)

```

```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yjbb.
GN YJBB OR STM4189.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Iatireille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 492-543 FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=94110223; PubMed=8282893;
RA Conlin C.A., Hakansson K., Liljas A., Miller C.G.;
RT "Cloning and nucleotide sequence of the cyclic AMP receptor protein-
RT regulated Salmonella typhimurium pepB gene and crystallization of its
RT product, an alpha-aspartyl dipeptidase."
RL J. Bacteriol. 176:166-172(1994).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO B.SUBTILIS YQEW.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008896; AAL23013.1; -.
DR EMBL; U01246; -; NOT_ANNOTATED_CDS.
DR StyGene; SG10492; yjbb.
DR InterPro; IPR003841; Na/Pi cotranspt.
DR InterPro; IPR004633; Napi cotransptII.
DR Pfam; PF02690; NaPi_cotrans; 1.
DR TIGRFAMs; TIGR01013; 2a58; 1.
DR TIGRFAMs; TIGR00704; Napi cotrn rel; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
SQ SEQUENCE 543 AA; 59444 MW; 434D9F724DFFEB67 CRC64;

Query Match 4.0%; Score 144.5; DB 1; Length 543;
Best Local Similarity 20.2%; Pred. No. 0.014;
Matches 89; Conservative 53; Mismatches 134; Indels 165; Gaps 13;

Qy 100 LILLIGFLYFFVCSLDLSAFLQVGGKAGQFSSNSINPNLLGLVIGLVTVLVQSS 159
Db 4 LHLLSAVALLVWGTHTVGTGMVFGARLTVLSR-SVEKTP-LAFCAGIGTALVQSS 61
Qy 160 STSTSVVSMVSSSLTVAAPITWGANIGTSTNTIVAMQVGRSEFRAPAGATVH 219
Db 62 NATTLVTSFVAQDLVALPALVILGADVGTALMARILTF----- 102
Qy 220 DFFNWLVLVLPVEVATHYLEITQLIVESHFKNGEDAPDLLKVTYKPTKLIQVLDK 279

```

Db 103 -DLWSLPLI-----FIGVIFLGR 122
 Qy 280 KVISQIANDEKAKNKLWIKWTKFTNKTQINTVSTANCSPSLCWTGCIQNTWTKN 339
 Db 123 K-:::--QSRAGLGRV-----133
 Qy 340 VTYKENIAKQHIFVNFHLPDLAVGTILLILSLVLOGCLIMVKILGVLKGVQVATVIK 399
 Db 134 -----GTGLLILALEL-----IVQVFTITQANGVQVI- 163
 Qy 400 KTINTDPPFPFAMLTG--YLAIVGAGMTFTVQSSSVFTSALTPLIGICVITIRAYPLT 457
 Db 164 -----FASLTGDMLDALIGAMPALISYSSLAVALLTATLTAAGILISFPVALCLV 213
 Qy 458 LGSNIGTITTTAILAALASPGNALSLOIALCHFFNFISGILHLWPI--PPTRL--PIRM 513
 Db 214 IGANLGSLLAWLNNSSAANAARR-----VALGSULFLKILGSLVILFPVHPLANIMDELSEL 269
 Qy 514 AKGLGNISAKYRWFVAFYLI 534
 Db 270 PK-----SELVIYHFVYNLV 285
 RESULT 9
 YD13 SCHPO STANDARD; PRT; 630 AA.
 AC Q92341,
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Hypothetical protein ClF8.03c in chromosome I.
 GN SPAC1F8.03C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks J., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Ezer P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880 (2002).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE YCL70C/YHL047C/YKR106W FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z81312; CAB03597.1; -.
 DR PIR; T38110; T38110.
 DR GenADB SPombe; SPAC1F8.03c; -.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 80 100 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 303 323 POTENTIAL.
 FT TRANSMEM 336 356 POTENTIAL.
 FT TRANSMEM 371 391 POTENTIAL.
 FT TRANSMEM 407 427 POTENTIAL.
 FT TRANSMEM 440 460 POTENTIAL.
 FT TRANSMEM 466 486 POTENTIAL.
 FT TRANSMEM 503 523 POTENTIAL.
 FT TRANSMEM 575 595 POTENTIAL.
 SQ SEQUENCE 630 AA; 70557 MW; 1072680A64539FE7 CRC64;
 Query Match 3.9%; Score 139.5; DB 1; Length 630;
 Best Local Similarity 20.6%; Pred. No. 0.037;
 Matches 140; Conservative 94; Mismatches 217; Indels 227; Gaps 33;
 Qy 6 ELGDAQENPDKYLEGAGQOQTPADPKSKETNKNTEAPVTKIPLPSYSTATLIDETEV 65
 Db 15 ELQDAKEEKSNGNFVPEKAFSDDEKSGSYTNE-----TVS-----KM 55
 Qy 66 DDPNWLTLQDSGIKWE--RDT-----KGKILCFQIGRL-----ILLGLFYFF-- 110
 Db 56 DN-----SLQRGVSKIEAVRDSYQNRKGYLAAYAGIAILACSWASAIQSSTYSYQV 110
 Qy 111 -----VCSLDILSAPQVGVGGAGQFSSNSIMNPL--LGLVIGLVTVLV 156
 Db 111 YATASFNRVTSITLBIATATLAISSVCKPIGK--SDITSRPMTYTLVLFYVIGIVV 167
 Qy 157 QSSSTSTSVVSSVSSSLTVAALPIIMGANIGTSTITNTIVALMQVGRSEPR-RAFAG 215
 Db 168 ASSTISAYVIGSVFISI-----GSSGLDYMLTV-----VGLTSLKWRGF-- 209
 Qy 216 ATVHDFNWLVLVLPVAVHVLITQLIVES-FHFANGEDADPLLKVITPFTKLI 274
 Db 210 -----NTALLSTPYATVWFTGFIVQGIIDSNRWGCGYGFALIMPAVMT-PAVIL 259
 Qy 275 VOLD-----KKVISQIAMNDEKAKNK--SLVKIW-----301
 Db 260 MYLBRQANKDENIKLIINY--CTEKNKNKQSKQKWLKAVLEVDLPGILLGVGWSILL 317
 Qy 302 -----C-KFTTNKTOINTVTP 316
 Db 318 LPFSLTSYAKNGKNSPMLAMVVGVIILAYSZYEMFIAPYPSCPKRWARTETITAVII 377
 Qy 317 STANCTSPSL-----CWTDDGQWTKMNTYKEN---TAKCHLFWNFHLPDLAVGTIL 367
 Db 378 DFFYLAGYLOSMYFTTYWILYDWSYEDTYFNNTWTALC--VGVVF-----AGAMH 429
 Qy 368 LTLISLLVLCGLINIVKILGSLVKQVATVIKKTINFDPPFPFAMLTGYLAILVAGAMTF 427
 Db 430 RVFHRVYKVLQIIGLVKIVG-----YCLIRNFATGKVDLW---SLILGMGGSF 479
 Qy 428 IVQSSSVFTSALTPLIGIGVITIERAVPLTLGSGNIGTMTTALALASP-----G 477
 Db 480 SVVGSOVSQASVPHQDLAI-----ASSLLPYTNIG---GAIGAAIASPISNKPXYLR 532
 Qy 478 NALRSSLIQALCHFFNIGSILLWYPIPTPLPIRWAKGLGNISAKYRWFVAFYLIIPFF 537
 Db 533 EYLPSSINDQTVNFYDSSLIREFV--GTEIRD-----GAIKAYSR-----SMFFL 578

Qy 538 LIP---ITVFLSLAGWR 552
 Db 579 LVPAVSLSFELAAFWQ 596

Best Local Similarity 19.3%; Pred. No. 0.2;
 Matches 122; Conservative 97; Mismatches 242; Indels 172; Gaps 26;

Qy 90 ILCPFOGIGELILL-----LGLYFFVCS-----LDLSSAFQVGGMACQFF 133
 Db 71 LVSVISGALITASYIGVTSFFLVYRCKPKIRKSLIIDLISGIPSVIFGLFASQIL 130
 Qy 134 SNSSIMSNPLGLVIGLVITVLOVSSSTSIYVSVSSSLLTVAALPIIMGANIGTSI 193
 Db 131 ---SIFPFDILKPLSLNLNVIAMLSNPIPIVISLITVNTVYVNDI-LSVVVSLGENK 186
 Qy 194 TMTVALMQVDRSE-----FRFAPAGATVDFNWLVLVLLPVEVATHYLEITQ 245
 Db 187 TSAIYKIIEKPKQLTILTAFAAISETMAVNF-----VLQSVNYQEVINN 235
 Qy 246 -----LIVESPEP-KNGEDA-----PDLK 264
 Db 236 NRFTSLDKTSGVSTFISNGDEQINGVLYIFGIIILVLSLNFPAISANPKTLE 295
 Qy 265 VTEPFTKLIQDKKVISQAMN-----DEKAKKSLVKIWKCTFTNKTQINVTPEST 318
 Db 296 --RYPFLKLSNFIYQVWFIPNIALFVDLTSTQSVKKI-----KVN----- 338
 Qy 319 ANCTSPSLCWDGION--WTMKQVITYKENIAKCHIFVNFHPLDPAVGTILLISLWLC 376
 Db 339 -NINERSLPEKRLQSVWIKLVFLK-----IFQELICTFLAFGVLAILLFVFIN 389
 Qy 377 GCLTMIVKILGSLV-----KQVATVKKITINTDPPFPA-----MLTGY-LAILV 421
 Db 390 GS--VAIINNNGSTVFSFADSTGRALVNTLIVITITITFPLALLIALLNRYNNSKV 447
 Qy 422 GAGMTFIVQSSSVFTSALTPLIGIVITIERAYPLTIGNIGTITTA-----ILAA 472
 Db 448 KNVFNVFIDLSMPSIYYGLFGLSPFL--RVQLSAGGANGTSLIAGLITISVILLFL 505
 Qy 473 LASGNALRS-SLQIALCHFFPNIS-----GILLWYPIPTPL-----PI 511
 Db 506 IRTCCQALNNVSDLRISAFALGISKREVIKTVLPSALKGLIVALILSINRIIETAPF 565
 Qy 512 RMAGLGNISAKYRFAVFLYLIIPFLIPLTVRG--LSLAGWRVLVGVGVVPIIILVL 569
 Db 566 FITSLGSSN-----LPHSLGPGQITTRIYGQIFISNSAISVLETSLVSVVFLIL 618
 Qy 570 CIRLLQRCPRVLPKLQWNPFLNMRSLKPW 602
 Db 619 LIFFSVLPSLFLNKKQVWIKSKFQSKLM 651

RESULT 11

OCN3 HUMAN STANDARD; PRT; 556 AA.
 ID OCN3 HUMAN STANDARD; PRT; 556 AA.
 AC 075751; Q9UP02;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DB Organic cation transporter 3 (Extraneuronal monoamine transporter)
 DE (EMT) (Solute carrier family 22, member 3).
 GN SLC22A3 OR EMT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE-Kidney;
 RX MEDLINE=921254; PubMed=10196521;
 RA Gruendemann D., Schechinger B., Rappold G.A., Schoemig E.;
 RT "Molecular identification of the corticosterone-sensitive
 extraneuronal catecholamine transporter.";
 RL Nat. Neurosci. 1:349-351 (1998).
 RN [2]
 RP SEQUENCE OF 513-556 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=99134299; PubMed=9933568;

RESULT 10

PSTA MYCGE STANDARD; PRT; 654 AA.
 AC P47651;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphate transport system permease protein pstA homolog.
 GN PSTA OR MG411.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Smal K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403 (1995).
 RN [2]
 RP SEQUENCE OF 136-211 FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing.";
 RL J. Bacteriol. 175:7918-7930 (1993).
 CC -1- FUNCTION: COULD BE PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM FOR PHOSPHATE; PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION
 CC OF THE SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PERMEASE FAMILY. CYSTW SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC ENBL; U39723; AAC71639.1; --
 CC ENBL; U01746; AAD10559.1; --
 CC PIR; E64245; E64245.
 CC TIGR; MG411; --
 CC InterPro: IPR000515; BPD transp.
 CC Pfam: PF00528; BPD transp; 2.
 CC PROSITE: PS00402; BPD TRANSP INN MEMBER; FALSE NEG.
 KW Transpote; Phosphate transport; Transmembrane; Complete proteome.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 113 133 POTENTIAL.
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 368 388 POTENTIAL.
 FT TRANSMEM 418 438 POTENTIAL.
 FT TRANSMEM 454 474 POTENTIAL.
 FT TRANSMEM 486 506 POTENTIAL.
 FT TRANSMEM 535 555 POTENTIAL.
 FT TRANSMEM 613 633 POTENTIAL.
 SQ SEQUENCE 654 AA; 73525 MW; 997ED093F7A907B4 CRC64;

Query Match

3.6%; Score 129; DB 1; Length 654;

RA Verhaegh S., Schweifer N., Barlow D.P., Zwart R.;
 RT "Cloning of the mouse and human solute carrier 22a3 (SLC22a3/SLC22a3)
 RT identifies a conserved cluster of three organic cation transporters on
 RT mouse chromosome 17 and human 6q26-q27.";
 RL Genomics 55:209-218(1999).
 RN [3]
 RP FUNCTION AND TISSUE SPECIFICITY.
 RC TISSUE=Kidney;
 RX MEDLINE=20425388; PubMed=10966924;
 RA Wu X., Huang W., Ganapathy M.E., Wang H., Kekuda R., Conway S.J.,
 RA Leibach F.H., Ganapathy V.;
 RT "Structure, function, and regional distribution of the organic cation
 RT transporter OCT3 in the kidney.";
 RL Am. J. Physiol. 279:F449-F458(2000).
 CC -!- FUNCTION: Mediates potential-dependent transport of a variety of
 CC organic cations. May play a significant role in the disposition of
 CC cationic neurotransmitters and neurotransmitters in the brain.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA, SKELETAL MUSCLE,
 CC PROSTATE, AORTA, LIVER, FETAL LUNG, SALIVARY GLAND, ADRENAL GLAND,
 CC KIDNEY AND BRAIN CORTEX. NO EXPRESSION DETECTED IN SPLEEN.
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
 CC CATION SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ001417; CAA04751.1; -;
 DR EMBL; AF078749; AAD20977.1; -;
 DR Genbank; HGNC:10967; SLC22A3.
 DR MIM; 604842; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005624; C:membrane fraction; TAS.
 DR GO; GO:0015101; F:organic cation transporter activity; TAS.
 DR GO; GO:0015695; P:organic cation transport; TAS.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR004749; Orgcat transp.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005829; Sug transporter.
 DR Pfam; PF00083; sugar tr; 1.
 DR TIGRfam; TIGR00838; 2A0119; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
 KW Transport; Ion transport; Transmembrane; Glycoprotein.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 182 202 POTENTIAL.
 FT TRANSMEM 241 261 POTENTIAL.
 FT TRANSMEM 269 289 POTENTIAL.
 FT TRANSMEM 381 401 POTENTIAL.
 FT TRANSMEM 468 488 POTENTIAL.
 FT TRANSMEM 498 518 POTENTIAL.
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 556 AA; 61279 MW; C3CA2D7DD21C658 CRC64;
 Query Match 3.6%; Score 128.5; DB 1; Length 556;
 Best Local Similarity 20.5%; Pred. No. 0.18;
 Matches 99; Conservative 63; Mismatches 173; Indels 147; Gaps 22;
 QY 111 VCSLILSSAPOLVGKMGAGOFFSSIMGNPLGLGVITVVLVSSSTSTSIIVSNV 170
 DB 160 ILNLGSLTGAFTL--GYAADRY---GRVIVILSLCLGVG-----TGVVAF 202
 QY 171 SSSLTTRRAIPLINGA-NIGTSTNTVITVALMGVDREFRFRAGATVHDF----- 222
 DB 203 PN--FPFVIFRFLQGVFGKGTWMTCTYVITVIVGSK---QRRIVGIVIQMFTLGIIL 257

QY 223 -----NW--LSVLVLPVEVATHYLEITQ-----LIVESPHKNGEDAPDLKLVTK 268
 DB 258 PGIAVFPNQGIGLAITLPLSFLFLYYVVPSPRWLITR-----KKGDKALQILIRIAK 313
 QY 269 PFTKLIVQDKKVISQIAMNDEKAKNKSIVKIMCKTFTNKTQINVTVPSTANCISPSLCW 328
 DB 314 CNGKYL-----SSNYSBITVDEEVSNPSFLDL-----VETPMRCKTILMP- 356
 QY 329 TDGIONWTKNVTYKNIKACQHFVNPHLPDLAVGTILLILSLVLCGLIMVILUGS 388
 DB 357 -----AMFTSAVYVQGLVWELGIIGNLYIDFFISGVVPELGLLIL-----LATERLGR 406
 QY 389 VLKGVQVAIVIKT---INTDFPFPFAMLTGYLALVCGAMTFTVQSSSVFTSALTPLIGI 445
 DB 407 RUPFAASINIVAGACIVTAPLPEGLAWRTVATLGRGLTMAFEIVILVNSEL----- 460
 QY 446 GVITIERAYPTLIGNSITGTTTITAILAALASPGNALRSSLOIALCHFFNFISGILWYPI 505
 DB 461 -----YPTTL-RNFGV-----SLCSGLCDF----- 479
 QY 506 FTRLPIMAKLGNISAKYRWAVFYLIFFFLIPLTVFGL--SLAGWVLY-----GVG 558
 DB 480 -----GGITAP---FLFLRLAAVWLELPLIFGILASICGGLHWMLPTEKGIA 524
 QY 559 VP 560
 DB 525 LP 526
 RESULT 12
 COX1 LOCOMI STANDARD; PRT; 512 AA.
 AC Q36421;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1).
 GN COI.
 OS Locusta migratoria (Migratory locust).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.
 CX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96139026; PubMed=8587138;
 RA Flook P.K., Rowell C.H.F., Gellissen G.;
 RT "The sequence, organization, and evolution of the Locusta migratoria
 RT mitochondrial genome."
 RL J. Mol. Evol. 41:928-941(1995).
 CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
 CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
 CC AND COPPER B.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS
 CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```
CC -----X80245; CAA56527.1; --
DR EMBL; X80245; T11467.
DR PIR; T11467.
DR HSSP; P00396; 20CC.
DR InterPro; IPR000883; COX1.
DR PFam; PF00115; COX1; 1.
DR PRINTS; PRO1165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;
KW Respiratory chain; Inner membrane.
FT METAL 60 IRON (HEME A AXIAL LIGAND) (PROBABLE).
FT METAL 239 COPPER B (PROBABLE).
FT METAL 243 COPPER B (PROBABLE).
FT METAL 243 COPPER B (PROBABLE).
FT METAL 289 COPPER B (PROBABLE).
FT METAL 290 COPPER B (PROBABLE).
FT METAL 375 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
FT METAL 377 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
FT METAL 377 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
FT METAL 377 IRON (HEME A3 AXIAL LIGAND) (PROBABLE).
FT CROSSLINK 239 1'-histidyl-3'-tyrosine (by similarity).
SQ SEQUENCE 512 AA; 56814 MW; 98031A229B00511C CRC64;

Query Match 3.6%; Score 128; DB 1; Length 512;
Best Local Similarity 20.1%; Pred. No. 0.18; Indels 274; Gaps 31;
Matches 122; Conservative 61; Mismatches 150;

QY 104 LGFLYFFVCSIDLSSAFOLVGKMGQFFENSSIMENPLGLVIGLVITVLVQSSSTST 163
DB 14 IGLTYEYF-----GAWAGMVGTSMTIRBELG-----QPGT 45
QY 164 SIIVSMVSSSLTTRAAIPT-----INGANIGTSINTVALMQVDRSFRAPAGATVHD 220
DB 46 MINDDOLYNIITAHAFVIFPMVMPIMIG-GFGNWLIVPLMIGAPDPAFPR-----MNN 98
QY 221 FFWLMLSVLLPYEVATHYLEITOLIVESPHFKNGSDA-----PDLKLVITKPTKL-- 273
DB 99 MSFW-----ILPSLT-----LLMSVVDVNGAGTGTWYTPPLASVIAHAGSVDL 144
QY 274 -IVOLDKKVISOI--AMDEKANKSLVKIWCKTFTNKTOINVTPSTANCTSPSLCWT 330
DB 145 AIFSILHAGVSSILGAIN-----FITTAIINRNNNTLDQT-----PLFVWS- 186
QY 331 GIQNTWKNVYKENTAKQHFVPHLPDLAVGTILLISLVLCGLIMVILKISVL 390
DB 187 -----VAITALLLLSLPLVLAGAITML----- 208
QY 391 KGQVATVTKTINTDRPPF-----FAWLTG-----YLAILVAG- 424
DB 209 -----LTDRLNLTSTFDAGGDPILYQHLFWFFGHPHYVILLPGGIIISHIVCOESG 262
QY 425 -----MTFVQSSVFTSALTPLIGIVGITIRAY----- 454
DB 263 KIESFGTIGTYAMLSTGLMGFTVWAHEWFT-----VGMVDVT--RAYFTSATMIIAVP 314
QY 455 -----PL--TLG-----SNIGTTTAAIALAASFGNALSSLOIA 487
DB 315 TGKIVFWMATLVGTQKFNPPLLWALGFIFLFTMGGLGLVIA-----NSSLDIV 365
QY 488 L-----CHFFFNIS-----GILLWYPIPTPLIRMAKGLGNISAKYRWFVAFVY 531
DB 366 LHDYTVVVAHSHVYLSNGAVFALMGIIQWYPL-FTGLTWNSS-----KWLKIQF 413
QY 532 LIIPFLPIPTVFG---LSLAG-----NRVLVGVGVV-----VFIIILVLC 570
DB 414 TIMFTIG-VNLTFFQHPGLAGMPRRISYDPDAVTSNNVSSIGSTISITGIIMPILIMW 472
QY 571 LKLQSR 577
DB 473 ESMIKQR 479

RESULT 13
LCCTP_BACHD
ID LCCTP_BACHD STANDARD; PRT; 524 AA.
AC Q9K529;
```

```
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE L-lactate permease.
GN LCCTP OR BH3936.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]_TaxID=86665;
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RA MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- FUNCTION: May play a role in L-lactate transport.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE LACTATE PERMEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF001520; BAB07655.1; -.
DR PIR; H84141; H84141.
DR InterPro; IPR003804; Lactate perm.
DR Pfam; PF02652; Lactate_perm; 1.
DR TIGRPFams; TIGR00795; LCCTP; 1.
KW Transport; Transmembrane; Complete proteome.
FT TRANSMEM 12 34 POTENTIAL.
FT TRANSMEM 38 60 POTENTIAL.
FT TRANSMEM 67 89 POTENTIAL.
FT TRANSMEM 127 149 POTENTIAL.
FT TRANSMEM 156 178 POTENTIAL.
FT TRANSMEM 193 215 POTENTIAL.
FT TRANSMEM 224 246 POTENTIAL.
FT TRANSMEM 250 267 POTENTIAL.
FT TRANSMEM 297 319 POTENTIAL.
FT TRANSMEM 339 361 POTENTIAL.
FT TRANSMEM 374 396 POTENTIAL.
FT TRANSMEM 411 433 POTENTIAL.
FT TRANSMEM 505 522 POTENTIAL.
SQ SEQUENCE 524 AA; 55628 MW; 0376ED4D9A31AFF6 CRC64;

Query Match 3.5%; Score 127.5; DB 1; Length 524;
Best Local Similarity 18.2%; Pred. No. 0.2;
Matches 83; Conservative 83; Mismatches 138; Indels 153; Gaps 21;

QY 87 KGKILCFQOIGR---HILLGLFLY-----FFVCSLDILSSAFQ----- 122
DB 102 RGSII-----GISDQRIQLLIGFCFNAFLGAGRGVPIAICAVLLVSLGFKPLQAML 157
QY 123 -LVGKMGAGOFFPNSSIMENPLGLVIGLVITVLVQSSSTSTSIIVSVSSSLITVRAAI 181
DB 158 CLIIANGASGAFGA-----IGIPVGIIDTLGLEQGVTSMDVSN-MTALTLPWINFTI 207
QY 182 PLINGANIGTSINTVALMQVDRSFRAPAGATVHDFFNWLVLVLLPEVATHYLE 241
DB 208 PFL-----LWLMD-----SWKGIKEILFALLIVTSVY 235
QY 242 IITQLIVESPHFKNGEDAPDLKLVITKPTKLIVQLDKK-----VISQIAMDDEKAK 293
DB 236 TVSQALLITIF-----IGPELADIIPSLLTWGLLALFLRWQPRNTFLNNGCSESHAS 289
QY 294 NKSIVKIWCKTFTNKTOINVTPSTANCTSPSLCWTGDGIQNTWMTKNVYKENTAK----- 348
```

Db 290 LKDVIKAW-----SPFYLLTFVFLWSUP--AFKGLLABGGALE 326
 Qy 349 -COHLEVFHLEPDL-----AVGTILLLSLLVLCGLINIVKILGSLVKQVATV 397
 Db 327 FAKWAFV---VFGSIEVGVDFGATGATALLAAVTTVTTKMRMK-----ESISL 375
 Qy 398 IKKTINTDPPPPFAW---LAGYLAILVAGMTF---IVQSSSVFTSALTPLIG-IGVI- 448
 Db 376 LKKVI-VDFSIPIMWICAILGIAKMTYGLTGMALGEAVATTGAPFFLSPILGWIGVFM 434
 Qy 449 -----TIRAVPLILGSGNIGTMTTALALASPG 477
 Db 435 TGSVVNNLTFAPIQTAGAIIGTNPISLVAANTAGG 471
 RESULT 14
 OCN3 MOUSE STANDARD; PRT; 551 AA.
 AC Q9WTW5; Q9R209;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Organic cation transporter 3 (Solute carrier family 22, member 3).
 GN SLC22A3 OR OCT3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 EN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC STRAIN=CD-1; TISSUE=Placenta;
 RX MEDLINE=99134299; PubMed=9933568;
 RA Verhaegh S., Schweifer N., Barlow D.P., Zwart R.;
 RT "Cloning of the mouse and human solute carrier 22a3 (SLC22a3)
 RT identifies a conserved cluster of three organic cation transporters
 RT on mouse chromosome 17 and human 6q26-q27.";
 RL Genomics 55:209-218 (1999).
 RN [2]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE=Kidney;
 RX MEDLINE=20425388; PubMed=10966924;
 RA Wu X., Huang W., Ganapathy M.E., Wang H., Kekuda R., Conway S.J.,
 RA Leibach F.H., Ganapathy V.;
 RT "Structure, function, and regional distribution of the organic cation
 RT transporter OCT3 in the kidney.";
 RL Am. J. Physiol. 279:F449-F458 (2000).
 CC -I- FUNCTION: Mediates potential-dependent transport of a variety of
 CC organic cations. May play a significant role in the disposition of
 CC cationic neurotoxins and neurotransmitters in the brain.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA AND IN KIDNEY
 CC CORTEX BUT NOT MEDULLA. LOW LEVELS FOUND IN BRAIN. IN THE KIDNEY,
 CC EXPRESSED SPECIFICALLY IN THE PROXIMAL AND DISTAL CONVOLUTED
 CC TUBULES AND WITHIN BOWMAN'S CAPSULE BUT NOT IN THE
 CC GLOMERULUS.
 CC -I- DEVELOPMENTAL STAGE: LEVELS ARE HIGH DURING GESTATION BUT DECREASE
 CC GREATLY TOWARDS THE END OF GESTATION.
 CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
 CC CATION SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF078750; RAD20978.1; -;
 DR EMBL; AF078748; RAD20238.1; -;
 DR EMBL; AF082566; RAD53007.1; -;
 DR MGD; MGI:133817; SLC22a3.
 DR InterPro; IPR007114; MFS.

DR InterPro; IPR004749; Orgcat transp.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005829; Sug transporter.
 DR InterPro; IPR003663; Sugar transp.
 DR Pfam; PF00083; sugar tr; 1.
 DR TIGRFAMs; TIGR00898; 2A0119; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
 KW Transport; ion transport; transmembrane; Glycoprotein.
 FT TRANSMEM 21 POTENTIAL.
 FT TRANSMEM 41
 FT TRANSMEM 177 197
 FT TRANSMEM 236 256 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 FT TRANSMEM 376 396 POTENTIAL.
 FT TRANSMEM 464 484 POTENTIAL.
 FT TRANSMEM 493 513 POTENTIAL.
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 551 AA; 61052 MW; C4D66AAC2EFCFDB CRC64;
 Query Match 3.5%; Score 127; DB 1; Length 551;
 Best Local Similarity 20.4%; Pred. No. 0.23;
 Matches 112; Conservative 73; Mismatches 205; Indels 160; Gaps 26;
 Qy 49 LIPSYSTAT-LIDEPTEVDPPNLFPLQDSGIKWSERDTKGKILCFQGIKRLILLOFL 107
 Db 94 LLEATNTSSELSCDPLTAFNRSAPLVSCSG--DWRVYETHSTIVSQFD----- 140
 Qy 108 YFFVCS---LDLSSAFOLVGKMGAGOFFSNSIMSNPLGLV---IGVLVTIVQSSS 160
 Db 141 --LUCSNAMLDITQALIN--GFLAGAF-----TLGVAADRYGRLLIYLISCFG 186
 Qy 161 TS-TSIVVSMVSS-SLLTVRAAIPIMGANITGTSITNTVIALMOVDSERFRAFAGATV 218
 Db 187 VGLTGVVAPAPNFVSFVIFRFLQGVFGK--GAWMTCTFVIVTEIVGSK--QRRIVGVI 241
 Qy 219 HDPFNWLSVLVLPVEVAT-HYLEITOLIVESHP-----KNGEDAP 260
 Db 242 QMFET-LGIILFGLIAYFTPSWQGLAISLPSFLIYYVVPSPRWLITRKQGEKAL 300
 Qy 261 DLKVIKPTKLLIVQLDKKVISQIAMDDEKAKNSLVKWKCTFTNKTQINVTVPSTAN 320
 Db 301 QILRRVAKNGKHL-----SSNYSFIVTDEVSNPSCLDL-----VTFPQWRK 344
 Qy 321 CTSPLCWTDGIONWTKVNTYKENIAKQHI FVNPHLPDLAVGTILLISLLVLCGLI 380
 Db 345 CTLILMF-----AWFTSAVYQGLVWRLIGLGNLYIDPFISGLVLPGLALLIL---- 393
 Qy 381 MIVKILGSLVKQVAVIKKT---INTDPPFPFPAWLTGVLATLVGNGMTFIVQSSSVFTS 437
 Db 394 LTIERLGRUFRFAASNI VAGSCLVTAFLPEGIPWLTVTATLGRIGITMAFEIVLVNS 453
 Qy 438 ALTPILGIVTIERAYPLTLGSGNIGTMTTALALASPCNALRSSLIQALCHFFFNISG 497
 Db 454 EL-----YPTTL-RNFGV-----SLCSGLCDF----- 474
 Qy 498 ILLWYPIPTLPIRWAKGLGNISAKYKWFVYLIIFPFILPLTVFGL--SLAGWRVLV 555
 Db 475 -----GGIAP---FLFRLLAAIWLELPLIIFGLASVCGGLVML 511
 Qy 556 -----GVGVP 560
 Db 512 LPEYKGLALP 521
 RESULT 15
 SNGI YEAST
 ID -SNGI YEAST STANDARD; PRT; 547 AA.
 AC P46950;
 DR 01-NOV-1995 (Rel. 32, Created)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: August 26, 2003, 15:47:17 ; Search time 118 Seconds
(without alignments)
1506.765 Million cell updates/sec

Title: US-10-052-664-1
Perfect score: 3597
Sequence: 1 MAPPELGDAQPNFKYLEG.....SREAQEVPAQSKTECTAL 689

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	3594	99.9	689	4	Q9POV7	Q9POV7 homo sapien
2	3570.5	99.3	690	4	Q95436	Q95436 homo sapien
3	3023	84.0	591	4	Q8WVA9	Q8WVA9 homo sapien
4	2926	81.3	695	11	Q9JJ09	Q9JJ09 rattus norv
5	2905	80.8	697	11	Q9DBP0	Q9DBP0 mus musculu
6	2899	80.6	697	11	Q9Z290	Q9Z290 mus musculu
7	2851	79.3	693	6	Q27960	Q27960 bos taurus
8	2098	58.3	397	4	Q8N2K2	Q8N2K2 homo sapien
9	2093	58.2	674	13	Q9PT82	Q9PT82 xenopus lae
10	2083	57.9	674	13	Q9PT83	Q9PT83 xenopus lae
11	2080	57.8	643	13	Q9DDR7	Q9DDR7 cyprinus ca
12	2058	57.2	636	13	Q91237	Q91237 pseudopleur
13	2047	56.9	632	13	Q9PDS6	Q9PDS6 brachydanio
14	1959.5	54.5	631	13	Q9PTQ8	Q9PTQ8 brachydanio
15	1717	47.7	637	11	Q9D2V6	Q9D2V6 mus musculu
16	1706.5	47.4	653	6	Q28361	Q28361 didelphis m

Q8N8W2 homo sapien
Q8N130 homo sapien
Q8K4R8 rattus norv
Q8V155 rattus norv
Q9D52 oncorhynchu
Q9D51 raja erinac
Q9D54 squalus aca
Q8C012 mus musculu
Q9D50 oncorhynchu
Q9DDR9 gallus gall
Q9W18 rattus norv
Q23530 caenorhabdi
Q9D55 squalus aca
Q9Z323 rattus norv
Q9D53 raja erinac
Q9DDR8 gallus gall
Q77701 bos taurus
Q87918 vibrio chol
Q8des1 vibrio vuln
Q8mcs3 corynebacte
Q9Z322 rattus norv
Q99X99 staphylococ
Q99YH9 streptococc
Q8NZW2 streptococc
Q8d4M4 vibrio vuln
Q8Y8R8 listeria mo
Q9S92 streptococc
Q8dQX9 streptococc

17 1696 47.2 632 4 Q8N8W2
18 1529 42.5 599 4 Q8N130
19 1475.5 41.0 601 11 Q8K4R8
20 1112 30.9 252 11 Q8V155
21 1046 29.1 290 13 Q9D52
22 1002 27.9 290 13 Q9D51
23 989 27.5 290 13 Q9D54
24 966.5 26.9 321 11 Q8C012
25 962 26.7 290 13 Q9D50
26 865 24.0 230 13 Q9DDR9
27 834 23.2 268 11 Q9W18
28 830 23.1 573 5 Q23530
29 814 22.6 230 13 Q9D55
30 773.5 21.5 337 11 Q9Z323
31 766 21.3 230 13 Q9D53
32 724 20.1 240 13 Q9DDR8
33 710.5 19.8 232 6 Q77701
34 559.5 15.6 382 16 Q87918
35 548.5 15.2 386 16 Q8des1
36 418.5 11.6 388 16 Q8mcs3
37 393.5 10.9 474 16 Q8FMC5
38 285.5 7.9 327 11 Q9Z322
39 283 7.9 555 16 Q99X99
40 280.5 7.8 543 16 Q99YH9
41 277.5 7.7 543 16 Q8NZW2
42 276.5 7.7 602 16 Q8d4M4
43 274.5 7.6 544 16 Q8Y8R8
44 263.5 7.3 543 16 Q9S92
45 262.5 7.3 543 16 Q8dQX9

ALIGNMENTS

RESULT 1

Q9POV7 PRELIMINARY; PRT; 689 AA.

ID Q9POV7
AC Q9POV7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sodium dependent phosphate transporter isoform Nap1-11b.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=20079161; PubMed=10610722;
RA Xu H., Bai L., Collins J.F., Ghishan F.K.;
RT "Molecular cloning, functional characterization, tissue distribution,
RT and chromosomal localization of a human, small intestinal sodium-
RT phosphate (Na+-Pi) transporter (SLC34A2).";
RL Genomics 62:281-284(1999).
DR EMBL; AF146796; AAF31328.1; -;
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR003841; Na/Pi_cotranspt.
DR Pfam; PF02690; Na_Pi_cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 1.
SQ SEQUENCE 689 AA; 75613 MW; 791CE76D36BD3F5 CRC64;

Query Match 99.9%; Score 3594; DB 4; Length 689;
Best Local Similarity 99.9%; Pred. No. 1.6e-252;
Matches 688; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPPELGDAQPNFKYLEGAGQOPTADPKSKETKNKNTAPVTKIELLPSTATLID 60
Db 1 MAPPELGDAQPNFKYLEGAGQOPTADPKSKETKNKNTAPVTKIELLPSTATLID 60
Qy 61 EPTEVDPPNWLTLQDSQIKWSRDTKGLICFFQIGRLILIGFLYFVCSLDLTSSA 120

DR EMBL; AF234239; AAL55657.1; JOINED.
 DR EMBL; AF234241; AAL55657.1; JOINED.
 DR EMBL; AF234242; AAL55657.1; JOINED.
 DR EMBL; AF234243; AAL55657.1; JOINED.
 DR EMBL; AF234244; AAL55657.1; JOINED.
 DR InterPro; IPR003841; Na/Pi cotranspt.
 DR Pfam; PF02690; Na/Pi cotrans; 2.
 DR TIGRFAMs; TIGR01013; 2a58; 1.
 FT NON_TER 591 591.
 SQ SEQUENCE 591 AA; 64789 MW; 32ACE5C2B5098A7 CRC64;
 Query Match 84.0%; Score 3023; DB 4; Length 591;
 Best Local Similarity 99.8%; Pred. No. 3.6e-211;
 Matches 590; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPWPELGDAQPNPDKYLEGAGQOPTAPDKSKETNNTEAPVTKIELLPSYSTATLID 60
 Db 1 MAPWPELGDAQPNPDKYLEGAGQOPTAPDKSKETNNTEAPVTKIELLPSYSTATLID 60
 QY 61 EPTVEDDPNLPDLQSGIKWSEDTKGKILCFQGIKRLILLGLFLYFFVCSLDILSSA 120
 Db 61 EPTVEDDPNLPDLQSGIKWSEDTKGKILCFQGIKRLILLGLFLYFFVCSLDILSSA 120
 QY 121 POLVGGKMAQGFSSNSIMSNPLLGVLVTLVQSSSTSTSTWSMVSSSLLTVRAA 180
 Db 121 POLVGGKMAQGFSSNSIMSNPLLGVLVTLVQSSSTSTSTWSMVSSSLLTVRAA 180
 QY 181 IPIIMGANIGTSTNTIVALMOVGDRSEFRAPAGATVHDFNNLSVLVLPVEVATHYL 240
 Db 181 IPIIMGANIGTSTNTIVALMOVGDRSEFRAPAGATVHDFNNLSVLVLPVEVATHYL 240
 QY 241 EITITOLIVSEFHFKNGEDAPDLKVTTPFKLIVOLDKKVISQIAMNDEKAKKSLVKI 300
 Db 241 EITITOLIVSEFHFKNGEDAPDLKVTTPFKLIVOLDKKVISQIAMNDEKAKKSLVKI 300
 QY 301 WCKTFTNKQINTVTVSTANTCPSLQWTDGIQNTWKNTVYKENTAKQHI FVNFHLP 360
 Db 301 WCKTFTNKQINTVTVSTANTCPSLQWTDGIQNTWKNTVYKENTAKQHI FVNFHLP 360
 QY 361 LAVGTILLILSLVLCGLIMIVKILGSLVKGQVATVIKKTINTDPPPPFAWLTGYLAI 420
 Db 361 LAVGTILLILSLVLCGLIMIVKILGSLVKGQVATVIKKTINTDPPPPFAWLTGYLAI 420
 QY 421 VGAGMTFIVQSSSVFTSALTPLIGIVTIERAYPLTIGSNIGTITTTAILAALASPGNAL 480
 Db 421 VGAGMTFIVQSSSVFTSALTPLIGIVTIERAYPLTIGSNIGTITTTAILAALASPGNAL 480
 QY 481 RSSIQIALCHFFNIGSILLWYPIPTPLPRMAKGLGNISAKYRWFAVYLIIFFFLI 540
 Db 481 RSSIQIALCHFFNIGSILLWYPIPTPLPRMAKGLGNISAKYRWFAVYLIIFFFLI 540
 QY 541 LTVFGLSLAGRWLVGVGVVVFIIILVLCRLQLQRCPRVLPKQLQNNWF 591
 Db 541 LTVFGLSLAGRWLVGVGVVVFIIILVLCRLQLQRCPRVLPKQLQNNWF 591
 RESULT 4
 Q9J09 PRELIMINARY; PRT; 695 AA.
 AC Q9J09;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Type IIB sodium-phosphate transporter.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Lung;
 RX MEDLINE=20340494; PubMed=10880371;
 RA Hashimoto M., Wang D.-Y., Kamo T., Zhu Y., Tsujiuchi T., Konishi Y.,

RA Tanaka M., Sugimura H.;
 RT "Isolation and Localization of Type IIB Na/Pi Cotransporter in the
 RT Developing Rat Lung.";
 RL Am. J. Pathol. 157:21-27 (2000).
 DR EMBL; AF157026; RAF76291.1; -.
 DR InterPro; IPR003841; Na/Pi cotranspt.
 DR Pfam; PF02690; Na/Pi cotrans; 2.
 DR TIGRFAMs; TIGR01013; 2a58; 1.
 SQ SEQUENCE 695 AA; 75992 MW; 3FAFE827527E0061 CRC64;
 Query Match 81.3%; Score 2926; DB 11; Length 695;
 Best Local Similarity 80.5%; Pred. No. 4.9e-204;
 Matches 545; Conservative 66; Mismatches 64; Indels 2; Gaps 2;
 QY 1 MAPWPELGDAQPNPDKYLEGAGQOPTAPDKSKETNNTEAPVTKIELLPSYSTATLI 59
 Db 1 MAPWPELGDAQPNPDKYLEGAGQOPTAPDKSKETNNTEAPVTKIELLPSYSTATLI 60
 QY 60 DEPTVEDDPNLPDLQSGIKWSEDTKGKILCFQGIKRLILLGLFLYFFVCSLDILSS 119
 Db 61 EEPFEGNDPDLQSGIKWSEDTKGKILCFQGIKRLILLGLFLYFFVCSLDILSS 120
 QY 120 AFOLVGGKMAQGFSSNSIMSNPLLGVLVTLVQSSSTSTSTWSMVSSSLLTVRA 179
 Db 121 AFOLVGGKMAQGFSSNSIMSNPLLGVLVTLVQSSSTSTSTWSMVSSSLLTVRA 180
 QY 180 APTIMGANIGTSTNTIVALMOVGDRSEFRAPAGATVHDFNNLSVLVLPVEVATHY 239
 Db 181 APTIMGANIGTSTNTIVALMOVGDRSEFRAPAGATVHDFNNLSVLVLPVEVATHY 240
 QY 240 LEITITOLIVSEFHFKNGEDAPDLKVTTPFKLIVOLDKKVISQIAMNDEKAKKSLVKI 299
 Db 241 LEKLTNLVLETFSPONGEDAPDLKVTTPFKLIVOLDKKVISQIAMNDEKAKKSLVKI 300
 QY 300 IWCKTFTNKQINTVTVSTANTCPSLQWTDGIQNTWKNTVYKENTAKQHI FVNFHLP 359
 Db 301 IWCKTFTNKQINTVTVSTANTCPSLQWTDGIQNTWKNTVYKENTAKQHI FVNFHLP 360
 QY 360 DLAVGTILLILSLVLCGLIMIVKILGSLVKGQVATVIKKTINTDPPPPFAWLTGYLAI 419
 Db 361 DLAVGTILLILSLVLCGLIMIVKILGSLVKGQVATVIKKTINTDPPPPFAWLTGYLAI 420
 QY 420 LVGAGMTFIVQSSSVFTSALTPLIGIVTIERAYPLTIGSNIGTITTTAILAALASPGNA 479
 Db 421 LVGAGMTFIVQSSSVFTSALTPLIGIVTIERAYPLTIGSNIGTITTTAILAALASPGNT 480
 QY 480 LRSLQIALCHFFNIGSILLWYPIPTPLPRMAKGLGNISAKYRWFAVYLIIFFFLI 539
 Db 481 LRSLQIALCHFFNIGSILLWYPIPTPLPRMAKGLGNISAKYRWFAVYLIIFFFLI 540
 QY 540 PLTVFGLSLAGRWLVGVGVVVFIIILVLCRLQLQRCPRVLPKQLQNNWF 599
 Db 541 PLTVFGLSLAGRWLVGVGVVVFIIILVLCRLQLQRCPRVLPKQLQNNWF 600
 QY 600 KPMDAVVSKFTGCFQMRCCCRVCRACCLLQCGPKCRCCKCEDLEAQQGQDVPVK 659
 Db 601 KPMDNIISLATSFCFORCCCRVCRVCCVCGC-KCCRCGCKCKNLEEBEKEQDVPVK 659
 QY 660 APETFDNITISERAOGE 676
 Db 660 ASGGFDNTAMSKEOQDE 676
 RESULT 5
 Q9DBP0 PRELIMINARY; PRT; 697 AA.
 ID Q9DBP0
 AC Q9DBP0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Solute carrier family 34 (Sodium phosphate), member 2.
 GN SLG34A2.
 OS Mus musculus (Mouse).


```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boirelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK004832; BAB23600.1; -.
DR MGD; MGI:1342284; Slc34a2.
DR InterPro; IPR003841; Na/Pi_cotransp.
DR PFam; PF02690; Na_Pi_cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
SQ SEQUENCE 697 AA; 76244 MW; 2A7B9384857BF16F CRC64;

Query Match 80.8%; Score 2905; DB 11; Length 697;
Best Local Similarity 80.0%; Pred. No. 1.6e-202;
Matches 543; Conservative 62; Mismatches 70; Indels 4; Gaps 3;

Qy 1 MAPPELGDAPNDPKYLEGAGQOPTAPDKSKETNK-NNTEAPVTKIELLPSPSTATLI 59
Db 1 MAPPELENAQPNFGKFTBEGASGPQSSIPAKDKASTNDNGTPVAKTELLPSYSLVLI 60
Qy 60 DEPTVEDDPAKLPPTLQDSGKWSRDTKGKILCFQFGIGRLILLGLYFFVCSLDLSS 119
Db 61 EBHPGTDPDLPELQDTGKWSRDTKGKILCFQGVGKFIILLGLYFLVCSLDVSS 120
Qy 120 AFQVGVGMAGQPFSSNSIMSNPLGLVIGLVITLVVQSSSTSTSIIVSMVSSLLTVRA 179
Db 121 AFQVGVGMAGQPFSSNSIMSNPVLAVIGLVITLVVQSSSTSSIIIVSMVASSLLTVRA 180
Qy 180 APTIMGANGTSTNTIVAMQVGRSEFRPAPAGATVHDFNWLVLVLPVEVATHY 239
Db 181 APTIMGANGTSTNTIVAMQVGRSEFRPAPAGATVHDFNWLVLVLPVEATHY 240
Qy 240 LEITQLIVSEHFKNGEDAPDLKLVTKPTKLIIVOLDKKVISOIAMNDEKAKNSLVK 299
Db 241 LEITLNLVLETFKNGEDAPDLKLVTKPTKLIIVOLDKKVISOIAMNDEKAKNSLVK 300
Qy 300 IWCKTFNKTQINTVTPSTANCTSPSLCWTGDIQNTWTKNTYKNIACQHIFFNPHLP 359
Db 301 IWCKSTINTEMNVTPSTDNCTSPSYCWTGDIQNTWTKNTYKNIACQHIFFNPHLP 360
Qy 360 DIAVGTILLILSLVLCGLIMIVKLGSLKGVATVTKTINTDPPPPFAWLTGYLAI 419
Db 361 DIAVGTILLILSLVLCGLIMIVKLGSLKGVATVTKTINTDPPPPFAWLTGYLAI 420
Qy 420 LVGAGMTFIVQSSVFTSALTPILIGVITIERAYPLTLGNSITGTTTTLAALASPNA 479
Db 421 LVGAGMTFIVQSSVFTSALTPILIGVITIERAYPLTLGNSITGTTTTLAALASPNT 480
Qy 480 LRSSIQIALCHFFPNISGILLWYPIPTFLPRMAKGLNISKAKYRFAVFLYIFPFLI 539
Db 481 LRSSIQIALCHFFPNISGILLWYPIPTFLPRMAKGLNISKAKYRFAVFLYIFPFLI 540

```

```

Db 361 DLAVGILLTVSLVZCSGLIMIVKLGSLRGQAVTIKTLNDFPPFPFAMLTGYLAI 420
QY 420 LVGAGMTFIVQSSSVFTSALTFLIGIVITIRAYPLTIGSNIGTITTAIALAASPNA 479
Db 421 LVGAGMTFIVQSSSVFTSALTFLIGIVISIRAYPLTIGSNIGTITTAIALAASPNT 480
QY 480 LRSLOIALCHFFNFISGILLWYIPPTIRPIRMAGLGNISAKYRWFVYLIIFFFLI 539
Db 481 LRSLOIALCHFFNFISGILLWYIPPTIRPIRLAKGLGNISAKYRWFVYLIIFFFVT 540
QY 540 PLTVFGLSLAGMEVLVGVVPIIILVLCILLOSRCPRVLPKLQWNFPLPWRSL 599
Db 541 PLTVFGLSLAGMEVLVGVVPIIILVLCILLOSRCPRVLPLKLEDFWNPFLPWRSL 600
QY 600 KPMDAVVSFTGCFQMRCCCRVCCACCLLGCPCCKCRCKCEDL--EEAQBQDVP 657
Db 601 KPMDNISLATTFCORRCCCRVCCVCCWVGC--KCRCKCRCKQGBEKEQDIP 659
QY 658 VKAPETFDNITISREAQGE 676
Db 660 VKASGAFDNAAMSKCODE 678

RESULT 7
Q27960
ID Q27960 PRELIMINARY; PRT; 693 AA.
AC Q27960;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sodium-dependent phosphate transporter.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95255303; PubMed=7737195;
RA Helps C.R., Murer H., McGiven J.D.;
RT "Cloning, sequence analysis and expression of the cDNA encoding a
RT sodium-dependent phosphate transporter from the bovine renal
RT epithelial cell line NBL-1."
RL Eur. J. Biochem. 228:927-930(1995).
DR EMBL; X81699; CAA57345.1; -.
DR InterPro; IPR003841; Na/Pi_cotransp.
DR Pfam; PF02690; Na_Pi_cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
SQ SEQUENCE 693 AA; 75826 MW; 6B1CC17FE5C13213 CRC64;

Query Match 79.3%; Score 2851; DB 6; Length 693;
Best Local Similarity 76.3%; Pred. No. 1.3e-198;
Matches 531; Conservative 82; Mismatches 73; Indels 10; Gaps 3;

QY 1 MAPPELGAQNPDKYLEGAAGQPTAPDKSKETKNNTEAPVTKIELPFSYSTATLID 60
Db 1 MAPPELENSQTSKTYTKADGQSAPKAPKATEKDTGPTPKIELVPSSTATLIE 60
QY 61 EPTVEDDWNLTQDSGKWSERDYGKILCFQGIKILLGLGFLYFFVCSIDLSSA 120
Db 61 EPTVEDDWNLTQDSGKWSERDYGKILCFQGIKILLGLGFLYFFVCSIDLSSA 120
QY 121 FQLVGGHAGQFFSNSSIMNSPLGLVIGLVVLVQSSSTSTSVISWSSSILLTVRA 180
Db 121 FQLVGGHAGQFFSNSSIMNSPLGLVIGLVVLVQSSSTSTSVISWSSSILLTVRA 180
QY 181 IPIIMGANIGTSITNTIVALMQGDRSEFRPAFAGATVHDFNWLVLVLLPVEVATHYL 240
Db 181 IPIIMGANIGTSITNTIVALMQGDRSEFRPAFAGATVHDFNWLVLVLLPVEVATHYL 240
QY 241 EITQIVTSFFHFKNGEDAPOLLKVIITKPTKLIIVOLDKKVISQIAMNDEKAKNSLVKI 300
Db 241 EITQIVTSFFHFKNGEDAPOLLKVIITKPTKLIIVOLDKKVISQIAMNDEKAKNSLVKI 300

```

```

Db 241 ERTNLVSEFFHFKNGEAPELLKAVITDPTFKLIILQDKSLINQIAMNDESVQKSKMIKI 300
QY 301 WKCTFNKTQINVTVPSTANTCPSLCWTDDGIONWTKNVTYKENIAKCOHIFWNFLHPD 360
Db 301 WKCTFNVTNRVNVVPSDENCPSLCWTDDGLYTWTKNVTYKENIAKCOHIFWNFLSD 360
QY 361 LAVGTILLLSLLVLCGLIMIVKILGSLRGQAVTIKTKINTDTPFPFPAWLTGYLAI 420
Db 361 AIVGTILLTSLLCTCLILVKKLGSLRGQAAVVIKTKINTDTPFPFPAWLTGYLAI 420
QY 421 VGAGMTFIVQSSSVFTSALTFLIGIVITIRAYPLTIGSNIGTITTAIALAASPNA 480
Db 421 VGAGMTFIVQSSSVFTSALTFLIGIVISIRAYPLTIGSNIGTITTAIALAASPNT 480
QY 481 RSLQIALCHFFNFISGILLWYIPPTIRPIRMAGLGNISAKYRWFVYLIIFFFLI 540
Db 481 KSLQIALCHFFNFISGILLWYIPPTIRPIRLAKGLGNISAKYRWFVYLIIFFFLI 540
QY 541 LTVFGLSLAGMEVLVGVVPIIILVLCILLOSRCPRVLPKLQWNFPLPWRSLK 600
Db 541 LAVFGLSLIGWVPLVGVASPIVLVILVVLVVKILQSPFCPSLPQKLSRWDPLPWRSLK 600
QY 601 PWDAAVSKFTGCFQMRCCCRVCCACCLLGCPCCKCRCKCEDL--EEAQBQDVPVKA 660
Db 601 PWDKLTSLTSCFQMRCCCRVCCACCLLGCPCCKCRCKCEDL--EEAQBQDVPVKA 660
QY 661 PETFDNITISREAQ----GEVPASDSK--TECTAL 689
Db 658 PEAFNNLMDKRAQDGVTKSEVDASGTVKIVSSVTAL 693

RESULT 8
Q8N2K2
ID Q8N2K2 PRELIMINARY; PRT; 397 AA.
AC Q8N2K2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ90534.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Ovarian carcinoma;
RA Isozaki T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075015; BAC11354.1; -.
DR InterPro; IPR001450; 4FE4S_ferredoxin.
DR TIGRFAMs; TIGR01013; 2a58; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 397 AA; 43726 MW; 23D42CBFC9555458 CRC64;

Query Match 58.3%; Score 2098; DB 4; Length 397;
Best Local Similarity 98.3%; Pred. No. 3.2e-144;
Matches 396; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 287 MNDEKAKNSLVKICKTFTNTKTIQINVTVPSTANTCPSLCWTDDGIONWTKNVTYKENI 346
Db 1 MNDEKAKNSLVKICKTFTNTKTIQINVTVPSTANTCPSLCWTDDGIONWTKNVTYKENI 346
QY 347 AKCOHIFWNFLHPDLAVGTILLLSLLVLCGLIMIVKILGSLRGQAVTIKTKINTDTP 406
Db 61 AKCOHIFWNFLHPDLAVGTILLLSLLVLCGLIMIVKILGSLRGQAVTIKTKINTDTP 406

```

```

QY 407 PFFAWLTYGAILVAGMTFVQSSSVFTSALTPLIGVITIERAYPLTLGSGTIT 466
DB 121 PFFAWLTYGAILVAGMTFVQSSSVFTSALTPLIGVITIERAYPLTLGSGTIT 180
QY 467 TAILAALASFGNALRSSLOIALCHFFPNISGILLWYPIPTPLPIRMAGLGNISAKYRW 526
DB 181 TAILAALASFGNALRSSLOIALCHFFPNISGILLWYPIPTPLPIRMAGLGNISAKYRW 240
QY 527 FAVFYLLIFELIPLTVFGLSLAGRWLVGVGVVVFIIILVCLLQSRCPVLPKLL 586
DB 241 FAVFYLLIFELIPLTVFGLSLAGRWLVGVGVVVFIIILVCLLQSRCPVLPKLL 300
QY 587 QNNWFLPLWMSLKPWDAVSKFTGCFQMRCCCCCRVCCACCLLGCPCCKCCKCED 646
DB 301 QN-----WMSLKPWDAVSKFTGCFQMRCCCCCRVCCACCLLGCPCCKCCKCED 354
QY 647 LEEAQEQDVPVKAPETFDNITISREAAQGEVPASDSKTECTAL 689
DB 355 LEEAQEQDVPVKAPETFDNITISREAAQGEVPASDSKTECTAL 397

RESULT 9
Q9PT82 PRELIMINARY; PRT; 674 AA.
AC Q9PT82;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAY-2003 (T-EMBLrel. 23, Last annotation update)
DE Sodium phosphate cotransporter.
GN XNAPI-B.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine; PubMed=9094212;
RX MEDLINE=97248089;
RA Ishizuya-Oka A., Stelow M.A., Ueda S., Shi Y.B.;
RT "Temporal and spatial expression of an intestinal Na+/PO4 3-
cotransporter correlates with epithelial transformation during thyroid
RT hormone-dependent frog metamorphosis.";
RL Dev. Genet. 20:53-66(1997).
DR EMBL: L78836; AAF21135.1;
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000005; HTHAc.
DR Pfam; PF02690; Na/Pi_cotransp.
DR TIGRFAMs; TIGR01013; 2a58; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY 1; 1.
SQ SEQUENCE 674 AA; 73837 MW; 70CE7520C98E73F1 CRC64;

Query Match 58.2%; Score 2093; DB 13; Length 674;
Best Local Similarity 61.5%; Pred. No. 1.4e-143;
Matches 428; Conservative 90; Mismatches 134; Indels 44; Gaps 13;

QY 1 MAPWPELGADQPNPKYLEG-----AAGQOFTADPKSKETKNKNTNTEAPVTKIELLPYST 55
DB 1 MPPFPEIDNFGFTGVDYDDSKFVMTGTFNPNVND-----GTPSDPEK-ELSTYST 51
QY 56 ATL-----IDPEVDDPNLPTLQDSIKWSEDTKGLICFFOGIGRLILGLFYFVC 112
DB 52 LSLCKETPEPEV-DPMWDEPKSTGPKWAEWTKQRLSVLGLWVKSVLVLLTYLFFVC 110
QY 113 SLDILSSAFQLVGGKAGOFFNSIMSINPLLIGVGLVTLVQSSSTSTSIIVSMVSS 172
DB 111 SLDILSSAFQLVGGKAGDIFKNHSLVSNPVAGIVGLVTLVQSSSTSTSIIVSMVSS 170
QY 173 SLLTVRAAIPITMANGTSTINTIVALMQVDSERFRAPAGATVHDFNWLVLVLLP 232
DB 171 NILTVRSAPITMANGTSTINTIVALMQSDRNEFRAPAGATVHDFNWLVLVLLP 230

```

```

QY 233 VEVATHLEIITQILIVSRFFKNGCEDAPLLKVIKTPFKLIVOLDKKVISQIANDDEKA 292
DB 231 IEVATGFLHLSALVKSSENIQTQDAPDMLKVIITEPLTKGIIQOLDKKVIQDIANGDPLA 290
QY 293 KNSLVKIMCKTFNKKQIINVVPSTANTCTSPSLCWTGQIONWTMKNVTYKNI-----AK 348
DB 291 QNKSLLKKEC---SYKTUWISLPSAENCTASLQWTD-----NNVTWTEGFEITKYK 341
QY 349 CQHIFVNFHPLDLAGVTLLILSLVLCGLIMIVKILGSLVKGOVATVIKKTINTDPPF 408
DB 342 CSHIFASTWPLDLAGVLLILSLVLCGLIMIVKILGSLVKGOVATVIKKTINTDPPF 401
QY 409 PFAWLYGAILVAGMTFVQSSSVFTSALTPLIGVITIERAYPLTLGSGTITTTA 468
DB 402 PFWLGYLAMLVAGMTFVQSSSVFTSALTPLIGVITIERAYPLTLGSGTITTTA 461
QY 469 ILAALASPGNALRSSLOIALCHFFPNISGILLWYPIPTPLPIRMAGLGNISAKYRWFA 528
DB 462 LLAALASPGETLQNSVQIALCHFFPNISGILLWYPIPTPLPIRMAGLGNISAKYRWFA 521
QY 529 VFYLLIFELIPLTVFGLSLAGRWLVGVGVVVFIIILVCLLQSRCPVLPKLLQW 588
DB 522 VYLLILCFILMPLVGLSVAGQALVGVAVPVTMIVAVISVNVVLQSKWPRILPDLKD 581
QY 589 WNFPLWMSLKPWDA-VVSKFTGCFQMRCCCCCRVCCACCLLGCPCCKCCKCEDL 647
DB 582 WDFLPKWMHSLKPWDA-CMLGASLWCKQF-CGCCGCKHCKGC-----KCC---KCHDK 630
QY 648 EEAQEQDVPVKAPETFDNITISREAAQGEVPASDSK 683
DB 631 EDEECDIETKPOALEHNDV-IDLSDEIKKPESEDEQ 665

RESULT 10
Q9PT83 PRELIMINARY; PRT; 674 AA.
AC Q9PT83;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAY-2003 (T-EMBLrel. 23, Last annotation update)
DE Sodium phosphate cotransporter.
GN XNAPI-A.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine; PubMed=9094212;
RX MEDLINE=97248089;
RA Ishizuya-Oka A., Stelow M.A., Ueda S., Shi Y.B.;
RT "Temporal and spatial expression of an intestinal Na+/PO4 3-
cotransporter correlates with epithelial transformation during thyroid
RT hormone-dependent frog metamorphosis.";
RL Dev. Genet. 20:53-66(1997).
DR EMBL: L78835; AAF21134.1;
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000005; HTHAc.
DR InterPro; IPR003841; Na/Pi_cotransp.
DR Pfam; PF02690; Na/Pi_cotrans; 2.
DR TIGRFAMs; TIGR01013; 2a58; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY 1; 1.
SQ SEQUENCE 674 AA; 73818 MW; E34B7809A346B5F4 CRC64;

Query Match 57.9%; Score 2083; DB 13; Length 674;
Best Local Similarity 61.1%; Pred. No. 7.3e-143;
Matches 425; Conservative 93; Mismatches 134; Indels 44; Gaps 13;

QY 1 MAPWPELGADQPNPKYLEG-----AAGQOFTADPKSKETKNKNTNTEAPVTKIELLPYST 55

```


Db 331 -----MEKCNHIFVDTGTPDLAVGLLLAGSLVWLCTCLILAVKMLNSLK 376
Qy 392 GOVATVKKTTINTDPPFPFALWLTGYLAILVAGMTFIVQSSSVFTSALTPLIGIVITIE 451
Db 377 GOVANVIQKINTDPPAPFTWVGYFAMVVGASMTFVQSSSVFTSAITPLIGLVISIE 436
Qy 452 RAYPLTLGSGNIGTTTTAILAALASPCNALRSSLOIALCHFFENISGILLWYPIPTRLPI 511
Db 437 RAYPLTLGSGNIGTTTTAILAALASPREKLSSSFQALCHFFENISGILLWYPIPTRLPI 496
Qy 512 RMAKGLGNISAKYRWFVYLIIFFFELIPLTVFGLSLAGKVLVGVGVFVFIILVLCL 571
Db 497 RMAKALGKETAKYRWFVYLIIVLCVLLPLSLVFGISMAGQAMVGVCTPFGALLAFVIV 556
Qy 572 RLQSRCPVLPKLIQNNWFLPLWMSLXPDWAVSKFTGCF 613
Db 557 NVLQSRSPGHLPKWLQTWDFLPRWMSLQPLDGLITRATLCY 598

Search completed: August 26, 2003, 15:52:43
Job time : 123 secs